

PF 29-SEP-1999; 99WO-US22976.
 XX
 ER 30-SEP-1998; 98US-0163648.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Acton LS, Robison KE, Hsieh FY;
 XX
 DR WPI: 2000-293140/25.
 DR N-PSDB; AAA12764.
 XX
 PT Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)
 PT polypeptide useful for detecting an ACE-2 therapeutic for treating
 PT hypertension, congestive heart failure, myocardial infarction,
 PT atherosclerosis and renal failure -
 XX
 PS Claim 2; Fig 1; 138pp: English.
 XX
 CC The present sequence represents a human angiotensin converting enzyme-2
 CC (ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The
 CC sequence of the full length ACE-2 cDNA was determined from a clone
 CC obtained from a cDNA library prepared from mRNA of a human heart of
 CC a subject who had congestive heart failure. ACE-2 has significant
 CC sequence homologies with ACE enzymes, and has also been shown to
 CC hydrolyse angiotensin I into Ang.(1-9). The ACE-2 therapeutics are
 CC used to treat blood pressure related diseases and conditions, such as
 CC hypertension, congestive heart failure, chronic heart failure, acute
 CC heart failure, myocardial infarction, atherosclerosis and renal
 CC failure.
 CC
 SQ Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 21; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 60
 DB 19 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 78
 QY 61 LAQWYPLQEIOWLVKQLQALQONGSSVLSSEDSKRLNTLNTMSTIYSTGKVCNDNP 120
 DB 79 LAQWYPLQEIOWLVKQLQALQONGSSVLSSEDSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLYEEYVVLKEMARAHYED 180
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLYEEYVVLKEMARAHYED 198
 QY 181 YGDIYRGDYEVNGVDYDYSKGLIEDVEHTEFEIKPLYEHLHAYVRATLMAAYPSYISP 240
 DB 199 YGDIYRGDYEVNGVDYDYSKGLIEDVEHTEFEIKPLYEHLHAYVRATLMAAYPSYISP 258
 QY 241 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEKEKFVS 300
 DB 259 IGCPLPAHLGDMGGRWTNLYSLTVPFGOKPNIDVDAMVDQAMQRIFFKEKEKFVS 318
 QY 301 GLPNMTQGEWENSMLTDPGNVQKAVCHPTAMDIGKGFRLIMCTKVTMDPFLTAHHEMGH 360
 DB 319 GLPNMTQGEWENSMLTDPGNVQKAVCHPTAMDIGKGFRLIMCTKVTMDPFLTAHHEMGH 378
 QY 361 IQYDMAVAAQPELLRNGANGGEHVAEGLMSTLSAAPKHLKSTGLSPDQEDNEEIN 420
 DB 379 IQYDMAVAAQPELLRNGANGGEHVAEGLMSTLSAAPKHLKSTGLSPDQEDNEEIN 438
 QY 421 LKQALITVGLTPEFYALEKRMWVFGKELPKDQMKKMWEMKREITVGVEVPYHDETVC 480
 DB 439 LKQALITVGLTPEFYALEKRMWVFGKELPKDQMKKMWEMKREITVGVEVPYHDETVC 498
 QY 481 DPASLFEVNSDYSEIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGQKLFNNL 540
 DB 499 DPASLFEVNSDYSEIRYRTLYQFOQALCOAANHEGFLAKCDISNSTEAGQKLFNNL 558
 QY 541 RLKSEFWTLALENVGAKMMNVRPLLTNTEPELFTMLKQDNKNSFVGWSTDMSPYADQSI 600

DB 559 RLKSEFWTLALENVGAKMMNVRPLLTNTEPELFTMLKQDNKNSFVGWSTDMSPYADQSI 618
 QY 601 KVRISLSKALGDKAYENDNEMTLFRSSVYARQYELKYNOMILFGEEDYRVANILKPR 660
 DB 619 KVRISLSKALGDKAYENDNEMTLFRSSVYARQYELKYNOMILFGEEDYRVANILKPR 678
 QY 661 ISFNEFYTAARKNSDIIPTREVEKAIMRSSRINDAFRLDNLSEFLGIQPTLGPNOBP 720
 DB 679 ISFNEFYTAARKNSDIIPTREVEKAIMRSSRINDAFRLDNLSEFLGIQPTLGPNOBP 738
 RESULT 2
 AAY67310
 ID AAY67310 standard; Protein; 805 AA.
 XX
 AC AAY67310;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE Human MPROT15 amino acid sequence #1.
 XX
 KW MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;
 KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;
 KW Alzheimer's disease; cytokine.
 XX
 OS Homo sapiens.
 XX
 PN JP11318472-A.
 XX
 PD 24-NOV-1999.
 XX
 PF 22-JAN-1999; 99JP-0014949.
 XX
 PR 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX
 PA (SMITK) SMITHKLINE BEECHAM PLC.
 XX
 DR WPI: 2000-109268/10.
 DR N-PSDB; AA259465.
 XX
 PT MPROT15 polypeptide and MPROT15 polynucleotides - useful for the
 PT treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.
 XX
 PS Claim 1; Page 15; 22pp; Japanese.
 XX
 CC This is amino acid sequence #1 of human MPROT15. The MPROT15
 CC polynucleotide and polypeptide sequences can be used for the treatment of
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous
 CC denaturation, Alzheimer's disease and diseases related to the processing
 CC of peptide hormones and cytokines.
 CC
 SQ Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 21; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 60
 DB 19 STEEOAKTFDKFHEADLFYQSSLASMTNTNTEENYQNNNAGDKMSAFLEKQST 78
 QY 61 LAQWYPLQEIOWLVKQLQALQONGSSVLSSEDSKRLNTLNTMSTIYSTGKVCNDNP 120
 DB 79 LAQWYPLQEIOWLVKQLQALQONGSSVLSSEDSKRLNTLNTMSTIYSTGKVCNDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLYEEYVVLKEMARAHYED 180
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLYEEYVVLKEMARAHYED 198
 QY 181 YGDIYRGDYEVNGVDYDYSKGLIEDVEHTEFEIKPLYEHLHAYVRATLMAAYPSYISP 240

XX	(MILL-) MILLENNIUM PHARM INC.
PA	
XX	
PI	Action SL, Robison KE;
XX	
DR	WPI; 2001-210604/21.
XX	
DR	N-PSDB; AAD0275B.
XX	
PT	Novel genes encoding angiotensin converting enzyme-2 useful as
XX	antisenze or antigene agents for therapeutics, diagnostics and
PI	screening assays -
XX	
PS	Claim 33; Fig 1; 76pp; English.
XX	
CC	The present amino acid sequence is human angiotensin converting enzyme-2
CC	(ACE-2), also referred as peptidyl dipeptidase A (PC 3.4.15.1). Nucleic
CC	acid sequence encoding ACE-2 is useful as antisenze or antigene agents
CC	for sequence specific modulation of gene expression or in the analysis of
CC	single base pair mutations in the gene. Nucleic acid sequence encoding
CC	ACE-2 is useful in therapeutics, diagnostics and in screening assays.
CC	ACE-2 antagonist is used to treat hypertension or congestive heart
CC	failure (CHF). ACE agonist is used to reduce the inflammation and pain
CC	resulting from an insect sting or bite, which was accompanied by an
CC	injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-2
CC	protein levels for determining the disease or condition associated with
CC	an aberrant protein level.
XX	
XX	Sequence 805 AA;
XX	
Query Match	100.0%; Score 3869; DB 22; Length 805;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches	0; Indels 0; Caps 0

Db	1	SYTERQAKFEJLDRFHNEAEDLEYQSSLASMNYNNTITEENVQNNNNAGDKWSAFKQST	78
QY	61	LAQMPPLQJQIONLTYKLOLOALQOONGSSVSEDSKSRNTLNTMSTITYSGKVCNDNP	120
Db	79	LAQMPPLQJQIONLTYKLOLOALQOONGSSVSEDSKSRNTLNTMSTITYSGVCNDNP	138
QY	121	QSCLLLEPGLNETMANSIDYNERLWAMESWRSYEGKOLRPLYEYVLLKNEMARAHYED	180
Db	139	QSCLLLEPGLNETMANSIDYNERLWAMESWRSYEGKOLRPLYEYVLLKNEMARAHYED	198
QY	181	YGDYWRGDEYVNGVGDYDYSRGOLIEDEYHFEETIKLEYHLIAYYRAKLMAYPSYISD	240
Db	199	YGDYWRGDEYVNGVGDYDYSRGOLIEDEYHFEETIKLEYHLIAYYRAKLMAYPSYISD	258
QY	241	IGCLPAHLIGDMGCFWNTNLSLTVPFGQKNIDYTDAMVDQAMDQORLEKEAKFPYFV	300
Db	259	IGCLPAHLIGDMGCFWNTNLSLTVPFGQKNIDYTDAMVDQAMDQORLEKEAKFPYFV	318
QY	301	GLPNNTOGFEMWNSMLTDPGNVOKAVCPRTAMDJGKGFRTIMCTKYTMDFLTAHHNGH	360
Db	319	GLPNNTOGFEMWNSMLTDPGNVOKAVCPRTAMDJGKGFRTIMCTKYTMDFLTAHHNGH	378
QY	361	IQYMAAYAAQFLLRNGANGFHBANGELISLSAATPRKILSIGLSLSPFOEDNETEINF	420
Db	379	IQYMAAYAAQFLLRNGANGFHBANGELISLSAATPRKILSIGLSLSPFOEDNETEINF	438
QY	421	LKQALTYIGTLPTTYMLEKRMNMFVFGELIPKXQNMKKMMKMKRELIVGVGEVPRHDEYIC	480
Db	439	LKQALTYIGTLPTTYMLEKRMNMFVFGELIPKXQNMKKMMKMKRELIVGVGEVPRHDEYIC	498
QY	481	DPASLFHVSNDYSIRIRYRTYLTLOFQFOEALCOAANHEGRLHKDISNSTEAGOKLFNML	540
Db	499	DPASLFHVSNDYSIRIRYRTYLTLOFQFOEALCOAANHEGRLHKDISNSTEAGOKLFNML	558
QY	541	RLGSEBPWTALAEVNGAKMMNVRPLNTFEPLETLKXQNNNSPFGWSTMSPPADDSI	600
Db	559	RLGSEBPWTALAEVNGAKMMNVRPLNTFEPLETLKXQNNNSPFGWSTMSPPADDSI	618

QY 601 KVRISLKSALGDKAYENNENEMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 660
 DB 619 KVRISLKSALGDKAYENNENEMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 678
 QY 661 ISFNEFYAPKNVSDIIPRTVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 720
 DB 679 ISFNEFYAPKNVSDIIPRTVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 738
 RESULT 4
 AAB48095
 ID AAB48095 standard; Protein: 805 AA.
 AC AAB48095;
 XX 19-MAR-2001 (first entry)
 DE Human Zace2 protein.
 XX Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 XX zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 XX ventricular systolic dysfunction; renal impairment; heart failure;
 XX scleroderma renal crisis; atherosclerosis; anti-inflammatory; human;
 XX antiarthritis; bradykinin inactivator.
 OS Homo sapiens.
 XX WO200070032-A1.
 PN 23-NOV-2000.
 PD 03-MAY-2000; 2000MO-US11932.
 PF 13-MAY-1999; 9905-0311482
 PR 27-AUG-1999; 9905-0384706
 XX (ZYMO) ZYMOGENETICS INC.
 PA Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX WPI: 2001-025018/03.
 DR N-PSDB; AAC84366, AAC84367.
 XX Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX Example 1; Page 95-100; 125pp; English.
 PS The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents the human Zace2 protein.
 XX Sequence 805 AA;
 S0 Query Match 100.0%; Score 3869; DB 22; Length 805;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STIEBOAKTFLDKFHEAEADLFYQSSLASWNTNTTEENYQNNMAGDKASATLKQST 60
 DB 19 STIEBOAKTFLDKFHEAEADLFYQSSLASWNTNTTEENYQNNMAGDKASATLKQST 78
 QY 61 LKOWPBLDIOUWLYKLOALQOONGSSVLSBDRSKRLNTLNTSTYSGKVCNPNP 120
 DB 79 LKOWPBLDIOUWLYKLOALQOONGSSVLSBDRSKRLNTLNTSTYSGKVCNPNP 138
 QY 121 OECLLEPGLNEMANSIDYNERLWMSWRSSEVCKOLRPLYEEVYVAKNEMARANHED 180
 DB 139 OECLLEPGLNEMANSIDYNERLWMSWRSSEVCKOLRPLYEEVYVAKNEMARANHED 198
 QY 181 YGDYWRGDEVNGVQCYDYSRGQLEDEYHTEFEELKPLYEHHAAYRAKIMNAPSYISP 240
 DB 199 YGDYWRGDEVNGVQCYDYSRGQLEDEYHTEFEELKPLYEHHAAYRAKIMNAPSYISP 258
 QY 241 IGLPAHLGDMWGRFTNLYSLVFPFGQKPNIDVTDAMVDQAMDAQRIEAEKFEVSV 300
 DB 259 IGLPAHLGDMWGRFTNLYSLVFPFGQKPNIDVTDAMVDQAMDAQRIEAEKFEVSV 318
 QY 301 GLPMTQGFENSMULDPCGVQKAVCHPTAMDLGKDFRLMCTVYTMDFLTAHENGH 360
 DB 319 GLPMTQGFENSMULDPCGVQKAVCHPTAMDLGKDFRLMCTVYTMDFLTAHENGH 378
 QY 361 IYDMMAYAAQFLIRNGANEGFEAVGEIMLSAATPKHLKSLGILSPDOEDNETEINF 420
 DB 379 IYDMMAYAAQFLIRNGANEGFEAVGEIMLSAATPKHLKSLGILSPDOEDNETEINF 438
 QY 421 LKQALITVGLPTTYLLEKRWKMKGEIPKQNMKKWEMKRELVGVVEVPHEHYC 480
 DB 439 LKQALITVGLPTTYLLEKRWKMKGEIPKQNMKKWEMKRELVGVVEVPHEHYC 498
 QY 481 DPASLPHVSNDSYFIRYRTFLYQFOEALCOAAHHEGLHRCDISNSTEAGOKLFNML 540
 DB 499 DPASLPHVSNDSYFIRYRTFLYQFOEALCOAAHHEGLHRCDISNSTEAGOKLFNML 558
 QY 541 RLGRSEPTLALENVGAKMNNRPLNTEPEFLTKDKONKSPFGMSSTWSPYADOST 600
 DB 559 RLGRSEPTLALENVGAKMNNRPLNTEPEFLTKDKONKSPFGMSSTWSPYADOST 618
 QY 601 KVRISLKSALGDKAYENNENEMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 660
 DB 619 KVRISLKSALGDKAYENNENEMTLFRSSVAYAMROYFLKKNOMILFGEEDVAVANLKPR 678
 QY 661 ISFNEFYAPKNVSDIIPRTVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 720
 DB 679 ISFNEFYAPKNVSDIIPRTVEKAIMRSRINDAFLRNDLSLEFLGIOTPLGPPNQP 738
 RESULT 5
 AAU99701
 ID AAU99701 standard; Protein: 805 AA.
 AC AAU99701;
 XX 24-SEP-2002 (first entry)
 DE Human angiotensin converting enzyme-2 (ACE-2) protein.
 XX Human: angiotensin converting enzyme-2; ACE-2; body weight disorder;
 KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;
 KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;
 KW familial partial lipodystrophy; hypercholesterolemia; hyperlipidaemia;
 KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;
 KW neurodegenerative disorder; peptide hormone; cytokine processing;
 KW myocardial infarction; cardiomyopathy; inflammatory bowel disease;
 KW systemic inflammation response syndrome; polytrauma; pain; stroke;
 KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;
 KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;
 KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;

KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;
 KM adult respiratory distress syndrome; wound healing; appetite;
 XX body mass index.
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT 1..18
 FT 19..805
 FT /label= Mature_human_ACE_2_protein
 XX
 PN W0200239997-A2.
 XX
 PD 23-MAY-2002.
 XX
 PE 31-OCT-2001; 2001WO-0845703.
 XX
 PR 01-NOV-2000; 2000US-0704216
 PR 29-MAY-2001; 2001US-087038Z.
 PR 19-OCT-2001; 2001US-371741P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Acton ST, Ocean TD, Gould AE, Dales NA, Guan B, Brown JA;
 PI Patane M, Kadambi VJ, Solomon M, Stricker-Krongrad A;
 XX
 DR WPI; 2002-547572/58.
 DR N-PSDB; ABR87623.
 XX
 PT Treating body weight disorder and increasing muscle mass comprises
 PT administering angiotensin converting enzyme-2 modulating compound
 XX
 PS Example 5; Page 387-390; 395pp; English.
 XX
 CC The present invention describes a new method of treating a body weight
 CC disorder, increasing muscle mass and decreasing body fat by
 CC administration of angiotensin converting enzyme (ACE)-2 modulating
 CC compound. The invention can be used for treating body weight disorders,
 CC particularly obesity of at least grade 1, diabetes, atherosclerosis and
 CC a state associated with lipid metabolism. The method is used for treating
 CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,
 CC generalised partial lipodystrophy, familial partial lipodystrophy,
 CC hypercholesterolaemia, hyperlipidaemia, an aberrant metabolic rate,
 CC congestive heart failure, chronic heart failure, left ventricular
 CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease and Huntington's disease),
 CC diseases associated with peptide hormones or cytokine processing,
 CC myocardial infarction, cardiomyopathy, systemic inflammation response
 CC syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and
 CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis
 CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema
 CC following focal injury, diffuse axonal injury, stroke, reperfusion
 CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic
 CC disorders including asthma, adult respiratory distress syndrome, wound
 CC healing and scar formation. The invention decreases the appetite,
 CC increases muscle mass and decreases body fat of subject having body mass
 CC index of greater than 23 (preferably 24.9)kg/m². The present amino
 CC acid sequence represents the human ACE-2 protein of the invention.
 XX
 SQ Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 23; Length 805;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STEEGAKFTLDFKFNHEADLFYQSSIASMNTNTNTEENYONNANNAGDKMSAPLKEGST 60
 DB 19 STEEGAKFTLDFKFNHEADLFYQSSIASMNTNTNTEENYONNANNAGDKMSAPLKEGST 78
 QY 61 LAQMTPLQETIONTYKALQALQONGSVLSSEDKSRMLTNTWSTIYSTGKVCNPDPNP 120
 DB 79 LAQMTPLQETIONTYKALQALQONGSVLSSEDKSRMLTNTWSTIYSTGKVCNPDPNP 138

QY 121 QECILLPEPLNEMANSIDYNERLWAMESWSEVSGKOLRPLYEEVVLKNEMARNHED 180
 DB 139 QECILLPEPLNEMANSIDYNERLWAMESWSEVSGKOLRPLYEEVVLKNEMARNHED 198
 QY 181 YGDYWRGDEYVNGVDYDYSRGQLIEDVEHFEETKPLYLEHIAVRAKIMANYSYSTP 240
 DB 199 YGDYWRGDEYVNGVDYDYSRGQLIEDVEHFEETKPLYLEHIAVRAKIMANYSYSTP 258
 QY 241 IGCPLRAHLIGDMWGFRTNLSLTVPFGQKPNIDVTAMDQAMDQADRIKFEAEKFEVS 300
 DB 259 IGCPLRAHLIGDMWGFRTNLSLTVPFGQKPNIDVTAMDQAMDQADRIKFEAEKFEVS 318
 QY 301 GLPNMTQGFWENSMTLDPGNVOKAVCHPTAMDIGKDFRILMCTKVTMDDLTAHHENG 360
 DB 319 GLPNMTQGFWENSMTLDPGNVOKAVCHPTAMDIGKDFRILMCTKVTMDDLTAHHENG 378
 QY 361 IOYDMAVYAAQPLLRNGANEHGEHVAEIMSLSAATPKHLKSIGLLSPQEDNETEINF 420
 DB 379 IOYDMAVYAAQPLLRNGANEHGEHVAEIMSLSAATPKHLKSIGLLSPQEDNETEINF 438
 QY 421 LKQALITVGTLPFTYMLEKRWMPKGEIPRDQMKRWMEKREIVGVPEVPHDETYC 480
 DB 439 LKQALITVGTLPFTYMLEKRWMPKGEIPRDQMKRWMEKREIVGVPEVPHDETYC 498
 QY 481 DPASLFEHNSDYSFRTYTRTLTYPOFOALCOQAKHEGPHKCDISNSTAGOKLFNML 540
 DB 499 DPASLFEHNSDYSFRTYTRTLTYPOFOALCOQAKHEGPHKCDISNSTAGOKLFNML 558
 QY 541 RLGKSEPTLALENVYGAKNMVRPLNVEFPLFTWLDQKNSEFVGMSTDMSPYADQSI 600
 DB 559 RLGKSEPTLALENVYGAKNMVRPLNVEFPLFTWLDQKNSEFVGMSTDMSPYADQSI 618
 QY 601 KVISLKSALGDKAEYNDNEMKYLERSSVAAYAMRYFLKAVNOMLISEEYVRANKPR 660
 DB 619 KVISLKSALGDKAEYNDNEMKYLERSSVAAYAMRYFLKAVNOMLISEEYVRANKPR 678
 QY 661 ISFNEFYVAPKPVSDIIRTEVEKAIKMSRSRINDAFRLNDSLEFLGIOTPLPPNQP 720
 DB 679 ISFNEFYVAPKPVSDIIRTEVEKAIKMSRSRINDAFRLNDSLEFLGIOTPLPPNQP 738

RESULT 6
 AAE20353
 ID AAE20353 standard; Protein; 805 AA.
 XX
 AC AAE20353;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human ACE-2 full-length protein.
 XX
 KW Human; angiotensin converting enzyme-2; ACE-2; therapy; hypertension;
 KW peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;
 KW myocardial infarction; heart failure; arrhythmia; renal failure; gene;
 KW inflammation; fertility; enzyme; EC 3.4.15.1.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide
 FT /label= signal_peptide
 FT 1..18
 FT 19..805
 FT /note= "Mature ACE-2 protein"
 FT 19..740
 FT /note= "Extracellular domain"
 FT 374..378
 FT /note= "Zinc binding domain (ZBD)"
 FT 741..765
 FT /note= "Transmembrane domain"
 FT 766..805
 FT /note= "Cytoplasmic domain"
 FT Domain
 XX

PN W0200212471-A2.
 XX 14-FEB-2002.
 XX 09-AUG-2001; 2001MO-US25059.
 XX 09-AUG-2000; 2000US-0635501.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Action S, Robison KE, Hsieh FY;
 PI WPI; 2002-257481/30.
 DR N-PSDB; AAD32586.
 XX
 PT Isolated human polypeptide, known as angiotensin converting enzyme-2,
 PT useful for treating or preventing the development of an abnormal blood
 PT pressure or related diseases, e.g. hypertension, heart failure or
 PT myocardial infarction.
 PS
 PS Claim 2; Fig 1; 218pp; English.
 CC The invention relates to human angiotensin converting enzyme-2 (ACE-2)
 CC polypeptides and polynucleotides. ACE-2 is also known as peptidyl
 CC dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
 CC for treating or preventing the development of abnormal blood pressure
 CC and diseases or disorders associated with the protein in a subject. The
 CC diseases include hypertension, hypotension, congestive heart failure,
 CC chronic heart failure, acute heart failure, myocardial infarction,
 CC atherosclerosis, arrhythmia and renal failure. They are also useful
 CC for treating inflammatory conditions and diseases relating to fertility.
 CC The present sequence is human full-length ACE-2 protein.
 CC
 XX Sequence 805 AA:
 Query Match 100.0%; Score 3869; DB 23; Length 805;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STIEQAKTFLDKFNHEADLFYQSSLASMNNTNITEENVQNNMAGDKMSAFLEKQST 60
 DB 19 STIEQAKTFLDKFNHEADLFYQSSLASMNNTNITEENVQNNMAGDKMSAFLEKQST 78
 QY 61 LAQMPYLOEIONLVKIQLOALQONGSSVSEDSKSKLNTLNTMSTTSGKVCNDNP 120
 DB 79 LAQMPYLOEIONLVKIQLOALQONGSSVSEDSKSKLNTLNTMSTTSGKVCNDNP 138
 QY 121 QECLEPGLNEIMANSIDYNERLWAMESRSEVKQRLPYEEYVVLKKNEMARANYED 180
 DB 139 QECLEPGLNEIMANSIDYNERLWAMESRSEVKQRLPYEEYVVLKKNEMARANYED 198
 QY 181 YGDYWRGDEYNGVDYDSRGQLEDEVENTFEIKPLYEHLAYVRAKIMNAVSYTSP 240
 DB 199 YGDYWRGDEYNGVDYDSRGQLEDEVENTFEIKPLYEHLAYVRAKIMNAVSYTSP 258
 QY 241 IGCPRHLIDGMWGRFTNLTSLVPPFGQKPNIDVTAMDOADARIFREKAEKFFVS 300
 DB 259 IGCPRHLIDGMWGRFTNLTSLVPPFGQKPNIDVTAMDOADARIFREKAEKFFVS 318
 QY 301 GLPMTGCFWENSMUTDPCGNVOKAVCHPTANDLKGDFRILMCVKYVMDDFLTAAHBMGH 360
 DB 319 GLPMTGCFWENSMUTDPCGNVOKAVCHPTANDLKGDFRILMCVKYVMDDFLTAAHBMGH 378
 QY 361 IQYDMAVAAPFLLRNGANGEGFHEAVGEMSLSATPRKHLKSGILLSPDEQEDNETINF 420
 DB 379 IQYDMAVAAPFLLRNGANGEGFHEAVGEMSLSATPRKHLKSGILLSPDEQEDNETINF 438
 QY 421 LLKQALTVGLPLPYLMEKRWAVFGCEIPKDDMMKMKMKREIYGVAVPRVHDEYTC 480
 DB 439 LLKQALTVGLPLPYLMEKRWAVFGCEIPKDDMMKMKMKREIYGVAVPRVHDEYTC 498
 QY 481 DPASLFHVSNDYSFIRYTRTYLOFOFOEALCOAAKHGPRHAKDISNSTEAGOKLPMNL 540

DB 499 DPASLFHVSNDYSFIRYTRTYLOFOFOEALCOAAKHGPRHAKDISNSTEAGOKLPMNL 558
 QY 541 RLKSEPTWTLAENVGAKNNANVRPLNYFEPLFTWLDQKNSEFVGMSTDMSPYADQSI 600
 DB 559 RLKSEPTWTLAENVGAKNNANVRPLNYFEPLFTWLDQKNSEFVGMSTDMSPYADQSI 618
 QY 601 KYRISLSKALGKAVEMNDNEMYLPRSSVAYAMROYELKVNQNTLFEEDYRVAANKPR 660
 DB 619 KYRISLSKALGKAVEMNDNEMYLPRSSVAYAMROYELKVNQNTLFEEDYRVAANKPR 678
 QY 661 ISFNFFVAPKKNVSDIIPTEVEKARMSRSRINDAFRLNDSLEFLGIQPTLGPNNPP 720
 DB 679 ISFNFFVAPKKNVSDIIPTEVEKARMSRSRINDAFRLNDSLEFLGIQPTLGPNNPP 738
 RESULT 7
 AAU09092
 ID AAU09092 standard; Protein; 711 AA.
 XX
 AC AAU09092;
 DE 20-DEC-2001 (first entry)
 XX
 XX Novel human protein NHP #1.
 XX
 XX Human; novel human protein; NHP; antidiabetic; antirheumatic;
 XX antiallergic; cytosolic; antiarteriosclerotic; vilaratory;
 XX neuroprotective; nootropic; antiparkinsonian;
 XX anti-human immunodeficiency virus; antiallergic; vasotropic; cardiatic;
 XX hypotensive; anorectic; antihypertensive; neuroleptic; anticonvulsant;
 XX antihuman; immunosuppressive; cerebroprotective; antitubercial;
 XX antinflammatory; antibacterial; antiparasitic; thyromimetic;
 XX immunomodulator; antiseborrheic; dermatological; vasoconstrictor;
 XX gastrointestinal disorder; cardiovascular disorder; hypertension;
 XX coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 XX cachexia; male infertility; impotence; testicular cancer; lung tumor;
 XX hyperproliferative disorder; pulmonary system disorder;
 XX central nervous system disorder; bone disorder;
 XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 XX Huntington's disease; schizophrenia; mania; dementia; paraneoplasia;
 XX panic disorder; learning disability; amyotrophic lateral sclerosis;
 XX psychosis; autism; sleep disorder; immune system disorder;
 XX Hashimoto's thyroiditis; musculo-skeletal system disorders;
 XX multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 XX diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 XX acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 XX inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 XX wound healing; chromosome X.
 OS Homo sapiens.
 XX
 XX
 FH Key
 FT Domain
 FT 1..681
 FT /label= Extracellular_domain
 FT 48..55
 FT /label= Immunogenic-epitope
 FT 110..118
 FT /label= Immunogenic-epitope
 FT 136..146
 FT /label= Immunogenic-epitope
 FT 151..158
 FT /label= Immunogenic-epitope
 FT 219
 FT /label= OTHER
 FT /note= "OTHER" Any amino acid encoded by WSI
 FT 240
 FT /label= OTHER
 FT /note= "OTHER" Any amino acid encoded by RCC
 FT 499
 FT /label= OTHER
 FT /note= "OTHER" Any amino acid encoded by NTT
 FT 682..698
 FT Domain

PK 43-MAI-1999; 9905-0311482.

PR 27-AUG-1999; 99US-0384706.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX WPI: 2001-025018/03.
 DR N-PSDB; AAC84368, AAC84369;
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Claim 7; Page 104-109; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme zace2, zace2, an angiotensin-
 CC converting enzyme is a zinc metalloprotease that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC zace2 gene, and to detect aberrations associated with the zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while zace2 antagonists are used for inducing
 CC infertility. The present sequence represents the mouse zace2-5 protein.
 CC
 XX
 SQ Sequence 805 AA;
 Query Match 84.0%; Score 3251; DB 22; Length 805;
 Best Local Similarity 83.1%; Pred. No. 3.6e-279;
 Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

QY 481 DPASLEHNSDYSEFRYYRTLYQFOFQALCOAKHGBPLHKCDISNSTEAGOKLFNML 540
 DB 499 DPASLEHNSDYSEFRYYRTLYQFOFQALCOAKHGBPLHKCDISNSTEAGOKLFNML 558
 QY 541 RIGKSEPTWLALENVGAKNANVRPLNFEPLFTWLDQKNSEFVQSTWSPYADDSI 600
 DB 559 SIGNSEPTWLALENVGAKNANVRPLNFEPLFTWLDQKNSEFVQSTWSPYADDSI 618
 QY 601 KVRISLSKSLGDKAYEMNDNEMYLFRSSVAYAMROYELKVNOMLGEEDVRYANLKR 660
 DB 619 KVRISLSKSLGDKAYEMNDNEMYLFRSSVAYAMROYELKVNOMLGEEDVRYANLKR 678
 QY 661 ISFNEFVYAPKNVSDIIPTEVEKARMSRSRINDAFRLNDSLEFLIOPTLGPNNQP 720
 DB 679 VSFYEFVYAPKNVSDIIPTEVEKARMSRSRINDAFRLNDSLEFLIOPTLGPNNQP 738
 XX
 PS RESULT 9
 AAB48098
 ID AAB48098 standard; Protein; 805 AA.
 XX
 AC AAB48098;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Mouse Zace2-10 protein.
 XX
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloprotease; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
 KW antihypertic; bradykinin inactivator.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT Region 19..613
 FT /note="fragment specifically claimed for"
 FT Region 19..708
 FT /note="fragment specifically claimed for"
 FT Region 19..738
 FT /note="fragment specifically claimed for"
 FT Region 19..805
 FT /note="fragment specifically claimed for"
 FT Region 133..542
 FT /note="fragment specifically claimed for"
 FT Region 344..542
 FT /note="fragment specifically claimed for"
 FT Region 371..402
 FT /note="fragment specifically claimed for"
 XX
 PN WO200070032-A1.
 XX
 PD 23-NOV-2000.
 XX
 PE 03-MAY-2000; 2000MO-0511932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX
 DR WPI: 2001-025018/03.
 DR N-PSDB; AAC84370.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Claim 7; Page 113-118; 125pp; English.

CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.

XX Sequence 555 AA;

Query Match 74.9%; Score 2897; DB 22; Length 555;

Best Local Similarity 99.8%; Pred. No. 5,2e-248;

Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNAGDKWSAFLEKQST 60
 DB 19 STIEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNAGDKWSAFLEKQST 78
 QY 61 LAQMPFLOEIONLVKQLQALQOONGSSVSEDSKRLNTLNTMSTYTGKVCNPDNP 120
 DB 79 LAQMPFLOEIONLVKQLQALQOONGSSVSEDSKRLNTLNTMSTYTGKVCNPDNP 138
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKQRLPYEEYVLKNEMARAHYED 180
 DB 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKQRLPYEEYVLKNEMARAHYED 198
 QY 181 YGDIWRGDIYVNGVDGYSRGQLEDEVEHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 240
 DB 199 YGDIWRGDIYVNGVDGYSRGQLEDEVEHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 258
 QY 241 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKPNIDVTDAWDAQMDAQRIFKEAKFVSV 300
 DB 259 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKPNIDVTDAWDAQMDAQRIFKEAKFVSV 318
 QY 301 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYTMDFLTAAHHEMGH 360
 DB 319 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYTMDFLTAAHHEMGH 378
 QY 361 IQYDMAVAAOPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSTIGLSPPQEDNETEINF 420
 DB 379 IQYDMAVAAOPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSTIGLSPPQEDNETEINF 438
 QY 421 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPIKDDMMKKWEMKREIVGVVEPRPHDETVC 480
 DB 439 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPIKDDMMKKWEMKREIVGVVEPRPHDETVC 498
 QY 481 DPASLHVSNDSYFIYRTLYFOFOQALCOAANHEGRLKCDISNSTEAGOKL 536
 DB 499 DPASLHVSNDSYFIYRTLYFOFOQALCOAANHEGRLKCDISNSTEAGOKL 554

RESULT 11

ID AAY67311 standard; Protein; 480 AA.

XX AAY67311;

XX AAY67311;

XX 11-APR-2000 (first entry)

XX Human MPR015 amino acid sequence #2.

DE MPR015; treatment; hypertension; human; myocardial disease; apoplexy;

XX MPR015; treatment; hypertension; human; myocardial disease; apoplexy;

KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;
 KW Alzheimer's disease; cytokine.
 XX Homo sapiens.
 XX JP11318472-A.
 XX 24-NOV-1999.
 XX 22-JAN-1999; 99JP-0014949.
 XX 13-MAY-1998; 98GB-0010373.
 XX 18-AUG-1998; 98GB-0018009.
 XX (SMIK) SMITHKLINE BEECHAM PLC.
 XX WPI; 2000-109268/10.
 XX MPR015 polypeptide and MPR015 polynucleotides - useful for the
 XX treatment of hypertension, myocardial diseases, apoplexy, heart
 XX diseases, nervous denaturation, Alzheimer's disease etc.
 XX Claim 19; Page 20-21; 22pp; Japanese.

XX This is amino acid sequence #2 of human MPR015. The MPR015
 CC polynucleotide and polypeptide sequences can be used for the treatment of
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous
 CC denaturation, Alzheimer's disease and diseases related to the processing
 CC of peptide hormones and cytokines.

SQ Sequence 480 AA;

Query Match 64.7%; Score 2505; DB 21; Length 480;

Best Local Similarity 100.0%; Pred. No. 2,6e-213;

Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNAGDKWSAFLEKQST 60
 DB 18 STIEQAKTFLDKFNEHAEEDLFYQSSLASWNTNTTEENVQNNAGDKWSAFLEKQST 77
 QY 61 LAQMPFLOEIONLVKQLQALQOONGSSVSEDSKRLNTLNTMSTYTGKVCNPDNP 120
 DB 78 LAQMPFLOEIONLVKQLQALQOONGSSVSEDSKRLNTLNTMSTYTGKVCNPDNP 137
 QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKQRLPYEEYVLKNEMARAHYED 180
 DB 138 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGKQRLPYEEYVLKNEMARAHYED 197
 QY 181 YGDIWRGDIYVNGVDGYSRGQLEDEVEHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 240
 DB 198 YGDIWRGDIYVNGVDGYSRGQLEDEVEHTEFEIKPLYEHLAAVYRAKLMAVPSYISP 257
 QY 241 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKPNIDVTDAWDAQMDAQRIFKEAKFVSV 300
 DB 258 IGCPLPAHLGDMWGFEWNTNLSLVTPFGQKPNIDVTDAWDAQMDAQRIFKEAKFVSV 317
 QY 301 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYTMDFLTAAHHEMGH 360
 DB 318 GLPNMTQGFWEWSMLTDPGNVQKAVCHPTAMDLSKGFRLIMCTKYTMDFLTAAHHEMGH 377
 QY 361 IQYDMAVAAOPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSTIGLSPPQEDNETEINF 420
 DB 378 IQYDMAVAAOPFLLRNGANGEGHEAVGEIMSLSAATPKHLKSTIGLSPPQEDNETEINF 437
 QY 421 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPIKDDMMKKWEMKREIVGVVEPRPHDETVC 480
 DB 438 LKQALTIIVGTLPTFLYMLEKRWAVFGEIPIKDDMMKKWEMKREIVGVVEPRPHDETVC 498

RESULT 12

ID AAU09102 standard; Protein; 261 AA.

XX AAU09102

AC AAU09102;
 XX
 DT 20-DEC-2001 (first entry)
 XX
 DE Novel human protein NHP #11.
 XX
 KW Human: novel human protein; NHP; antidiabetic; antirheumatic;
 KW antiarthritic; cyostatic; antiarteriosclerotic; vulnerrary;
 KW neuroprotective; nootropic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antiallthmatic; vasotropic; cardiant;
 KW hypotensive; anorectic; antinfertility; neuroleptic; anticonvulsant;
 KW antimanic; immunosuppressive; cerebroprotective; antileucobial;
 KW antinflammatory; antibacterial; antipsoriatic; thyromimetic;
 KW immunomodulator; antiseborrheic; dermatological; vasoconstriction;
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;
 KW hyperproliferative disorder; pulmonary system disorder;
 KW central nervous system disorder; bone disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; schizophrenia; mania; dementia; paranoia;
 KW panic disorder; learning disability; amyotrophic lateral sclerosis;
 KW psychosis; autism; sleep disorder; immune system disorder;
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;
 KW multiple sclerosis; ischaemic brain injury; stroke; infectious disease;
 KW diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;
 KW acquired immunodeficient syndrome; leukaemia; Rheumatoid arthritis;
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 KW neural system disorder; respiratory disorder; olfactory disorder;
 KW wound healing.
 XX
 KW Homo sapiens.
 OS
 PN WO2001/4896-A1.
 PD
 XX 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US10542.
 XX
 PR 03-APR-2000; 2000US-194118P.
 PR 29-SEP-2000; 2000US-236384P.
 XX
 PA (HUMAN) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 DR WPI: 2001-626394/72.
 DR N-PSDB: AAS14890.
 XX
 PT New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders
 PS
 XX Claim 11: Page 311-312; 318pp; English.

CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, Rheumatoid arthritis, inflammatory Bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence represents an NHP of the invention.
 XX
 SQ Sequence 261 AA;
 Query Match 35.1%; Score 1359; DB 22; Length 261;
 Best Local Similarity 99.6%; Pred. No. 4.9e-112;
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 44 MNNAGDWSAFLEKQSTLAQMTPLQETQNTYKIQLOALQONSSVISEKSRRLTIIN 103
 DB 1 MNNAGDWSAFLEKQSTLAQMTPLQETQNTYKIQLOALQONSSVISEKSRRLTIIN 60
 OY 104 TMSITTSYGKVCNPDNPOECILLEPGINEIMANSLDNERLMMESRSRSGQLRPLYE 163
 DB 61 TMSITTSYGKVCNPDNPOECILLEPGINEIMANSLDNERLMMESRSRSGQLRPLYE 120
 OY 164 EYVLAKNEMARANHEDYGDYWRGDEYVNGVDYDSRGQULIEDVETFEIRPLYEHLH 223
 DB 121 EYVLAKNEMARANHEDYGDYWRGDEYVNGVDYDSRGQULIEDVETFEIRPLYEHLH 180
 OY 224 AYVAKLMNAPYSITISIGCLPAHLGDMMGRFTNLTYSITVPGQKPNIDVTDAVDOA 283
 DB 181 AYVAKLMNAPYSITISIGCLPAHLGDMMGRFTNLTYSITVPGQKPNIDVTDAVDOA 240
 OY 284 WDAQRIFREAEKEF 296
 DB 241 WDAQRIFREAEKEF 253
 RESULT 13
 ID AAR10426
 XX AAR10426 standard; Protein: 732 AA.
 AC AAR10426;
 XX
 DT 10-APR-1991 (first entry)
 XX
 DE Human testicular angiotensin conversion enzyme.
 XX
 KW human testicular angiotensin conversion enzyme; tACE;
 KW male sterility.
 XX
 OS Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT Peptide 1..21
 FT Protein /label- signal peptide 22..732
 FT /label- mature tACE
 XX
 PN W09100354-A.
 XX
 PD 10-JAN-1991.
 XX
 PF 05-JUL-1990; 90WO-FR00513.
 XX
 PR 05-JUL-1989; 89FR-0009062.
 XX
 PA (INNM) INST NAT SANTE RECH.
 XX
 PI Soubrlier F, Albenc-Gelas F, Hubert C, Corvol P;
 DR WPI: 1991-036748/05.
 DR N-PSDB: AA010328.
 XX
 PT Nucleic acid - encoding human testicular angiotensin conversion
 PT enzyme, used e.g. for in vitro detection of enzyme in organism
 XX

PS Claim 1; Fig 1; 48pp; French.

XX A bank of human testicular cDNA in Lambda gtl1 was screened with a
CC probe containing the final 3248 nucleotides of endothelial ACE. The
CC complete sequence of tACE was reconstructed from 4 separate clones.
CC The isolated nucleic acid sequence was inserted into a plasmid for
CC expression of the protein. The invention covers polypeptides
CC containing all or part of tACE sequence. These are useful in
CC treatment of inflammation or infectious diseases, especially acute
CC pancreatitis, or diseases in which kinins are involved. Antibodies
CC against the polypeptides are useful as immunoassay reagents for
CC tACE.

XX Sequence 732 AA:

Query Match 34.6%; Score 1337; DB 12; Length 732;
Best Local Similarity 41.7%; Pred. No. 2,2e-109;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEEQATFLDKFNEHEDLFYQSSLASMTNTNITE-----NQVNNNAGDKWSA 53
DB 70 TDEAASKFVEEYDPTQVWVNEAEANWNTNTTETSKILLQKNQJANHNT----- 123
QY 54 FLKEOSTLAOMYPLQEIONTLVKLOALQONGSSVSEDKSKRLNTLNTNSTYSGK 113
DB 124 --LKGTQARKRDVNOJLQNTTITKRIKKVODLERALPAQLEENKILDMETTSVAT 181
QY 114 VCNPDNPOECLELLEPGLNEIMANSIDYNERIMAMESRSEVGKQLRPLYEEYVVLKNEMA 173
DB 182 VCHPNG--SCIQLEPDLTNVMAATSRKYEDLLMAMEGRDAGRAIILGFYKYVELLNQAA 239
QY 174 RANHEDYGDYWRGDIYVNGVDGIDYSRGQLEDVEHTFEIKPLYEHLAIVAKILMAA 233
DB 240 RLMNGYVAGSGWRSMTYTPSLE-----ODLERLFOELPLYLNLAAVYRRLHNR 289
QY 234 Y-PSYISPIGCLPAHLGLDMGRRFTNLSTLVPEGQKPIIDTLDVAVDQAMQRIEKE 292
DB 290 YGAOHILEGRIPRAHLGLGNMAQTSNIYDLYVPFSPAPSDTTENALQSGWTPRRMFE 349
QY 293 AEKFEVSGLPNTQGFENSMILDPGNVOKAVCHPTAMDLAGK-DEKILMCTKVTWMDP 351
DB 350 ADDEFTSLGLPVPPEFMNKSMEKPTDGREVYCHASAMDFYNGKDERIKQCTVMEDEL 409
QY 352 LTAHNEGHIOYDMAAARPLRANGNEFHNAGEIMLSATKRLKSLGSLSPDRQ 411
DB 410 VVAHNEGHIOYDMQYDLRVALREGANPEFHAIQGVLAISTPRHLISLMLSEGG 469
QY 412 EDNETEINFLKQALITVGLTPTVLMKRRMVFGEIKPKDOMKKMWEKREIYGYE 471
DB 470 SD-EHDINFLMKALDKIAIFPSYLVQKRMVFDGISTIKENTNDEWMSLRKYGLCP 528
QY 472 PVPHEDEYCDPASLEHVSNDYSFIRYTRTLVQFOEALCQAAKHEGLAKDISNSTE 531
DB 529 PVPRTQGDFFPAKFNHSSVPIRYFVSFIIOFHEALCQAAHGTGLPKKDDIYSKE 588
QY 532 AGOKLEFMALGSEPTLALENVGAKNNANRPLANTPELTWIKDKQK--NSVYGV- 588
DB 589 AGORLATYAMKLGSRPPEAMOLITQPNMSASAMLSYKPLDMWTENELHGEKLGWP 648
QY 589 STDMSPYADQS 599
DB 649 QYNWTPMSARS 659

RESULT 14
AAO20501
ID AAO20501 standard; Protein; 1265 AA.

XX AAO20501;
XX 27-JUN-2002 (first entry)
XX Protein of APP related human homologue hcp51674.

XX Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;
KM APP pathway modulator; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 38 /label= unknown

XX MISC-difference 39 /note= "Encoded by TGN"

XX /label= unknown

XX /note= "Encoded by CNA"

XX W0200226820-A2.

XX 01-OCT-2001; 2001WO-EP11345.

XX 29-SEP-2000; 2000US-236893P.

XX 14-JUN-2001; 2001US-298309P.

XX (NOVS) NOVARTIS AG.

XX (NOVS) NOVARTIS-ERRINDONGEN VERW GES MBH.

XX Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;

XX Reinhardt MWM, Zusan S;

XX WPI; 2002-315796/35.

XX N-PSDB; AAK93935.

XX New transgenic fly, containing DNA encoding an Abeta portion of human

XX APP, useful for identifying agents which modulate the APP pathway and

XX which can be used to treat Alzheimer's disease -

XX Example 4; Page 94-97; 129pp; English.

XX The invention relates to a transgenic fly whose genome comprises DNA

XX encoding a polypeptide having the Abeta portion of human amyloid

XX precursor protein (APP), fused to a signal sequence. The DNA sequence

XX encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in

XX the specification. The DNA sequence is operably linked to a tissue-

XX specific expression control sequence. Expression of the sequence gives

XX the fly an altered phenotype. The purpose of the invention is for

XX identifying agents that inhibit or promote the expression and/or function

XX of genes or encoded polypeptides which modify the APP pathway. The agent

XX is a compound, triple helix DNA, antisense oligonucleotide, double

XX stranded RNA molecule, ribozyme, or particularly an antibody. It is used

XX to treat conditions such as Alzheimer's disease. The agent can be used as

XX an APP pathway modulator or in gene therapy. This sequence represents the

XX protein of the APP related human homologue hcp51674.

XX Sequence 1265 AA:

QY Query Match 34.6%; Score 1337; DB 23; Length 1265;
Best Local Similarity 41.7%; Pred. No. 5,2e-109;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEEQATFLDKFNEHEDLFYQSSLASMTNTNITE-----NQVNNNAGDKWSA 53
DB 603 TDEAASKFVEEYDPTQVWVNEAEANWNTNTTETSKILLQKNQJANHNT----- 656
QY 54 FLKEOSTLAOMYPLQEIONTLVKLOALQONGSSVSEDKSKRLNTLNTNSTYSGK 113
DB 657 --LKGTQARKRDVNOJLQNTTITKRIKKVODLERALPAQLEENKILDMETTSVAT 714
QY 114 VCNPDNPOECLELLEPGLNEIMANSIDYNERIMAMESRSEVGKQLRPLYEEYVVLKNEMA 173
DB 715 VCHPNG--SCIQLEPDLTNVMAATSRKYEDLLMAMEGRDAGRAIILGFYKYVELLNQAA 772
QY 174 RANHEDYGDYWRGDIYVNGVDGIDYSRGQLEDVEHTFEIKPLYEHLAIVAKILMAA 233


```

Db      773 KLNGYDADSDMSRSTETSTLE-----QDERLFOELQPLYLNLHAAYRALHRH 822
Qy      234 Y-PSYISPIGCLPAHLIDGMGRFTNLXSLTVPGOKENIDVTDAVDAQDAORIFEK 292
Db      823 YGAGHINLMBGPTRPAHLILGMMNAQWTSNIYDLVPPSPASMDTTEMLKOGTPRRMFE 882
Qy      293 AEKFFVSVGLPNNTOGFENSMILTDGNNVOKAVCHPTAMDLCG-DFTILMCTKVTMD 351
Db      883 ADDEFTSLGLLPVPEFENKSMLEKPTDREYVYASADDFNGKDFRIKCTTVNLDEL 942
Qy      352 LTAHHEMGIQYDAAVAAQPLLRNGANGCFHEAVGEIMSLSAAPKHLKSLGSLSPDQ 411
Db      943 VVAHHEMGIQYFMQYKDLVPAIRBAGANCFHEALIGVLAISVTPKHLHSLNLSSEGG 1002
Qy      412 ENNETEINFLKQALITVGLPTVYMLEKRMWVFKGEIPKQDMKKWMEKREIYVVE 471
Db      1003 SD-EHDINFLMKALDKIAIFPSYLVQWRWRFVDSITKKNYNQEWMSLAKYQGLCP 1061
Qy      472 PVPHDETCDPASLHVNDYSFIRYTYTLVQFOFQALCOAKHNEGPLKCDISNTE 531
Db      1062 PVPRTOGDDPFAKFIPISSVYIRYFVSFIIOFQFHEALCOAGHGTGLHKCDIYQSK 1121
Qy      532 ACQKLEFNLRLGSEPMTLALENVYGAKNMNYRPLNPEPLFTWIKQONK--NSFVGW- 588
Db      1122 ACQRLATANKLGFSPRWRPAMQITQPRMSASAMLSIFKPLLDWLTENELHGEKLGMP 1181
Qy      589 STDMSPYADQS 599
Db      1182 QYNWTPNSARS 1192

RESULT 15
AA04111
ID AA04111 standard; peptide; 1306 AA.
XX
AC AA04111;
XX
DT 07-SEP-1990 (first entry)
XX
DE Human angiotensin converting enzyme (ACE):
XX
KW human angiotensin converting enzyme; hypertension; bradykinin.
XX
OS synthetic.
XX
FH Key
FH Protein
FT Location/Qualifiers
FT 30..1277
FT /label=mature ACE
FT /note="derived from pre-ACE by removal of signal peptide"
FT Modified-site
FT 38..38
FT /label=putative N-glycosylation site
FT Modified-site
FT 54..56
FT /label=putative N-glycosylation site
FT Modified-site
FT 74..76
FT /label=putative N-glycosylation site
FT Modified-site
FT 111..113
FT /label=putative N-glycosylation site
FT Modified-site
FT 146..148
FT /label=putative N-glycosylation site
FT Modified-site
FT 160..162
FT /label=putative N-glycosylation site
FT Modified-site
FT 318..320
FT /label=putative N-glycosylation site
FT Modified-site
FT 445..447
FT /label=putative N-glycosylation site
FT Modified-site
FT 509..511
FT /label=putative N-glycosylation site
FT Modified-site
FT 523..525
FT /label=putative N-glycosylation site
FT Modified-site
FT 677..679
FT /label=putative N-glycosylation site
FT Modified-site
FT 713..715
FT /label=putative N-glycosylation site

```

```

FT      /label=putative N-glycosylation site
FT      760..762
FT      /label=putative N-glycosylation site
FT      942..944
FT      /label=putative N-glycosylation site
FT      1191..1193
FT      /label=putative N-glycosylation site
FT      1225..1227
FT      /label=putative N-glycosylation site
Db      WO9003435-A.
Pn
Pd      05-APR-1990.
Pf
Pe      27-SEP-1989; 89WO-FR00496.
Pr      27-SEP-1988; 88FR-0012620.
Pa      (INRM) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
Pi      Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P;
Pr      WPI: 1990-132272/17.
Dr      N-PSDB; AA004027.
XX
XX      Human angiotensin converting enzyme hydrolyses angiotensin I and kinins.
XX      Either intact enzyme or fragments thereof can be used to generate
XX      antibodies for diagnostic use. Oligonucleotide probes can also be made
XX      which are complementary to the sequence encoding the enzyme.
XX
XX      Disclosure: ; p; French.
XX
XX      New DNA encoding human angiotensin converting enzyme used eg in
XX      diagnosis of hypertension, evaluation of enzyme inhibitors
XX
XX      PS
XX
XX      Sequence 1306 AA:
SQ
Query Match 34.6%; Score 1337; DB 11; Length 1306;
Best Local Similarity 41.7%; Pred. No. 5.5e-109;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;
Qy      2 TIEQAKTFLDKFHEADLFYQSSLASWNTNITTE-----NVQNMNNAQDKWSA 53
Db      644 TDEAFASFEVHEYDRTQVWVNEFAEAMWNTNITTEKSKILLQKMOJANH----- 697
Qy      54 FLKEOSTLAQMYPLQEIQLNLTVKLQALQONSSVSEKSKRLNTILTMETIYSTGK 113
Db      698 --LKTGQARKFDVNOQNTTIKRIKYODLEALPAGDELYKTLMDETTYSVAT 755
Qy      114 VCNPDNPECILLEPGINEIMANSLDYNERLMAWESRSEVQGLRPLYEYVLKNEMA 173
Db      756 VCHPNG--SCIQLEPDLTNVATSRKEYEDILMAEGWRDAGRALLQGFYKVELLNQA 813
Qy      174 RANFVEYGDWRDGVNVCYDVGSRGQLEDEVEHTEPRDELRYEHLAAVRAKLMA 233
Db      814 KLNGYDADSDMSRSTETSTLE-----QDERLFOELQPLYLNLHAAYRALHRH 863
Qy      234 Y-PSYISPIGCLPAHLIDGMGRFTNLXSLTVPGOKENIDVTDAVDAQDAORIFEK 292
Db      864 YGAGHINLMBGPTRPAHLILGMMNAQWTSNIYDLVPPSPASMDTTEMLKOGTPRRMFE 923
Qy      293 AEKFFVSVGLPNNTOGFENSMILTDGNNVOKAVCHPTAMDLCG-DFTILMCTKVTMD 351
Db      924 ADDEFTSLGLLPVPEFENKSMLEKPTDREYVYASADDFNGKDFRIKCTTVNLDEL 983
Qy      352 LTAHHEMGIQYDAAVAAQPLLRNGANGCFHEAVGEIMSLSAAPKHLKSLGSLSPDQ 411
Db      984 VVAHHEMGIQYFMQYKDLVPAIRBAGANCFHEALIGVLAISVTPKHLHSLNLSSEGG 1043
Qy      412 ENNETEINFLKQALITVGLPTVYMLEKRMWVFKGEIPKQDMKKWMEKREIYVVE 471
Db      1044 SD-EHDINFLMKALDKIAIFPSYLVQWRWRFVDSITKKNYNQEWMSLAKYQGLCP 1102

```


KW vascular disorder; asbestosis.
 XX
 XX Mus sp..
 OS
 OS WO200136632-A2.
 PN 25-MAY-2001.
 PD
 XX 17-NOV-2000; 2000WO-IL00766.
 XX
 XX 17-NOV-1999; 99IL-0132978.
 PR 10-DEC-1999; 99IL-0133455.
 XX
 XX (COMP-) COMPUGEN LTD.
 PA
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI WPI; 2001-336004/35.
 DR N-PSDB; AAS06057.
 XX
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies
 XX
 XX Claim 4; Fig 57; 519pp; English.
 PS
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasocactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various
 CC disorders including cardiovascular diseases such as arteriosclerosis,
 CC myocardial infarction and coronary arterial thrombosis, renal diseases
 CC such as diabetic nephropathy, muscular diseases such as hypertrophy,
 CC immune disorders such as immune complex nephritis, multiple sclerosis,
 CC cancer, sarcoidosis, noncolidotic pulmonary granulomatous diseases such
 CC as asbestosis and vascular pathologies involving an endothelial
 CC abnormality such as deep vein thrombosis.
 CC
 CC Sequence 1249 AA;
 SQ
 Query Match 34.5%; Score 1334; DB 22; Length 1249;
 Best Local Similarity 42.6%; Pred. No. 9.5e-109;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;
 QY 2 TIEQAKTFLDKNEHAEDELYQSSIASWNTNTITEENVQNNANAGDKSAFLKQSTL 61
 DB 649 TDEAKADREVEEDRTAOVLINLEYAANMOYNTNITEGSKILLERSTEVSHHTLYGR 708
 QY 62 AOMYPLAEIONLTVKQLALQONGSSVLEDSKRLNTLNTMTSTYSGKCNPDNQ 121
 DB 709 AKTFDVSNFONSSIKRIKILQNDRAVLPKRELEYNOLLDMETTYISLNCYNG-- 766
 QY 122 ECLLEFGLNEIANSIDYNERLWAMESRSEVKOLRPLYEEVYVAKNEMARANEYD 181
 DB 767 TCMPLPEDLTNMATSRELLWAMKSRDKGRALIPFPKRYEVSNRKIALNGTTDA 826
 QY 182 GDVWRGDIYEVNGDGYISRGQILIEVDEHFEELKPYEHLHAYVRKALNNAPVS-YISP 240
 DB 827 GDSWRSLSYEDNLE-----QDEKLYOEOLPLYLMLHAYVRSLSLRHNGSEITN. 876
 QY 241 IGCPLAHLIDMMGRFTNLSLVTPGQKPNIDVDVAVQADQRIKFEKAEFEVSV 300
 DB 877 DGPFLAHLIDMMGRFTNLSLVTPGQKPNIDVDVAVQADQRIKFEKAEFEVSV 300
 QY 301 GLPMTGCEWENSMALDPCGNVQAVCHPTAMDJGK-DFTILMCTKTKTMDFLAHHEMG 359
 DB 937 GLPMTGCEWENSMALDPCGNVQAVCHPTAMDJGK-DFTILMCTKTKTMDFLAHHEMG 936

QY 360 HIOYDMAVAAPPELLNGANEGFHEAVGEIMSLAATPKHLKSLIGLLSPDQENDETEIN 419
 DB 997 HIOYFMOYDLPVTEFEGANPCFHEAIGDIMALSVTPKHLKSLINLSTE-GSGYEYDIN 1055
 QY 420 FLKQALITVGLTPFTYMEKRWKWFKEIIPKDDMKKKWEMKREIVGVPEPHDETY 479
 DB 1056 FLKQALITVGLTPFTYMEKRWKWFKEIIPKDDMKKKWEMKREIVGVPEPHDETY 479
 QY 480 CDPASLFHVSNDYSFIRYTRTYLPQFOFQALCOAKHSGPLKCDISNSTAGOKLFNM 539
 DB 1116 FDPGSKFHPANPYRATFSPFIQFQFHALCRAGHTGPKKCDIYQSKRAGKLADA 1175
 QY 540 LRLGSEPTLALENVGAKNANVRPLNFEPLFTWLDQNK--NSFVGW-STDMSP 594
 DB 1176 MKLGSKFHPANPYRATFSPFIQFQFHALCRAGHTGPKKCDIYQSKRAGKLADA 1233
 RESULT 18
 AA002985
 ID AA002985 standard; Protein; 1252 AA.
 XX
 XX AA002985;
 XX
 XX 12-SEP-2001 (first entry)
 DE
 XX Angiotensin converting enzyme (ACEV) splice variant protein #85.
 XX
 XX Angiotensin converting enzyme splice variant; ACEV; interleukin 6;
 KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
 KW platelet-derived endothelial cell growth factor; cardiovascular disease;
 KW cellular tumour antigen p53; cyclin-dependent kinase inhibitor 1C;
 KW vasocactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
 KW myocardial infarction; coronary arterial thrombosis; renal disease;
 KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
 KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
 KW noncolidotic pulmonary granulomatous disease; endothelial abnormality;
 KW vascular disorder; asbestosis.
 KW
 XX Mus sp..
 OS
 OS WO200136632-A2.
 PN 25-MAY-2001.
 PD
 XX 17-NOV-2000; 2000WO-IL00766.
 XX
 XX 17-NOV-1999; 99IL-0132978.
 PR 10-DEC-1999; 99IL-0133455.
 XX
 XX (COMP-) COMPUGEN LTD.
 PA
 XX Levine Z, David A, Azar I, Khosravi R, Bernstein J;
 PI WPI; 2001-336004/35.
 DR N-PSDB; AAS06085.
 XX
 XX Novel alternative splicing variants e.g. variant of angiotensin
 PT converting enzyme (ACEV), useful in identifying candidate compounds
 PT capable of binding to the variant and to detect anti-variant antibodies
 XX
 XX Claim 4; Fig 85; 519pp; English.
 PS
 XX The sequence represents an angiotensin converting enzyme splice variant
 CC (ACEV) polypeptide. The polypeptides of the invention include variants of
 CC granulocyte colony stimulating factor receptor, glucagon, interleukin 6,
 CC platelet-derived endothelial cell growth factor, cyclin-dependent kinase
 CC inhibitor 1C, cellular tumour antigen p53, and vasocactive intestinal
 CC polypeptide receptor 2. The polypeptides and their associated nucleic
 CC acids are useful for identification of variant sequences and detection of
 CC candidate compounds capable of binding the molecules. The sequences of
 CC the invention can be used in the treatment and diagnosis of various

disorders including cardiovascular diseases such as arteriosclerosis, myocardial infarction and coronary arterial thrombosis, renal diseases such as diabetic nephropathy, muscular diseases such as hypertrophy, immune disorders such as immune complex nephritis, multiple sclerosis, cancer, sarcoidosis, nonrheumatic pulmonary granulomatous diseases such as asbestosis and vascular pathologies involving an endothelial abnormality such as deep vein thrombosis.

Sequence 1252 AA:

Query Match 34.5%; Score 1334; DB 22; Length 1252;

Best Local Similarity 42.6%; Pred. No. 9, 5e-109;

Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 2 TIEBAKTYLDFKFNHEADLFYQSSIASWNTNTTEENYONMANNAGDMSAFLEKQSTL 61
 DB 649 TIEBAKTYLDFKFNHEADLFYQSSIASWNTNTTEENYONMANNAGDMSAFLEKQSTL 61
 QY 62 AOMYPLQETIONLTVLQALQONSSVLSDESKRLNTLNTMSTYSGKVCNPDNPQ 121
 DB 709 AKTFVSNFQNSSTIRIKLQNDRAVLPRKEEYVQILDMETYSLSNICYTNG-- 766
 QY 122 ECLLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPIYEEYVLKNEARAHNEDY 181
 DB 767 TCMPLPEPLTNMATSRYEELMAMKSMRDKVGRALIPFPKVEFSNIAKLNGTTDA 826
 QY 182 GDMYSGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 240
 DB 827 GDMYSGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 240
 QY 241 IGCPLPAILGDMGCFWNTLSLVPPGOKPNDIVTAMDOAMDORIKREKFEVSV 300
 DB 877 DGPPIAHLLGMMAGQTSNITDVAAPPSAPNIDATAMTKQGTPTRIKFEADNFTSL 936
 QY 301 GLPNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 DB 937 GLPNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 QY 937 GLPNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 DB 937 GLPNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 QY 360 HIOYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 DB 997 HIOYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 QY 420 FLKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 DB 1056 FLKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 QY 480 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 539
 DB 1116 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 539
 QY 540 LRLGSEPTWLALENVVGANNVRLPLNYEPLFTWMLDKON--NSFVGM-STDMSP 594
 DB 1176 LRLGSEPTWLALENVVGANNVRLPLNYEPLFTWMLDKON--NSFVGM-STDMSP 594
 QY 1176 MKLGSYKPEAKMLKLTGQPNMSASAMMYFKLTENLTENRKGELTGMPEYMAP 1233

RESULT 19

ABB64006 standard; Protein; 615 AA.

ABB64006;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 18810.

Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL08109.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

PS Disclosure; SEQ ID NO 18810; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572022).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

Sequence 615 AA;

Query Match 27.6%; Score 1069; DB 22; Length 615;
 Best Local Similarity 36.3%; Pred. No. 1e-85;
 Matches 217; Conservative 120; Mismatches 239; Indels 22; Gaps 9;

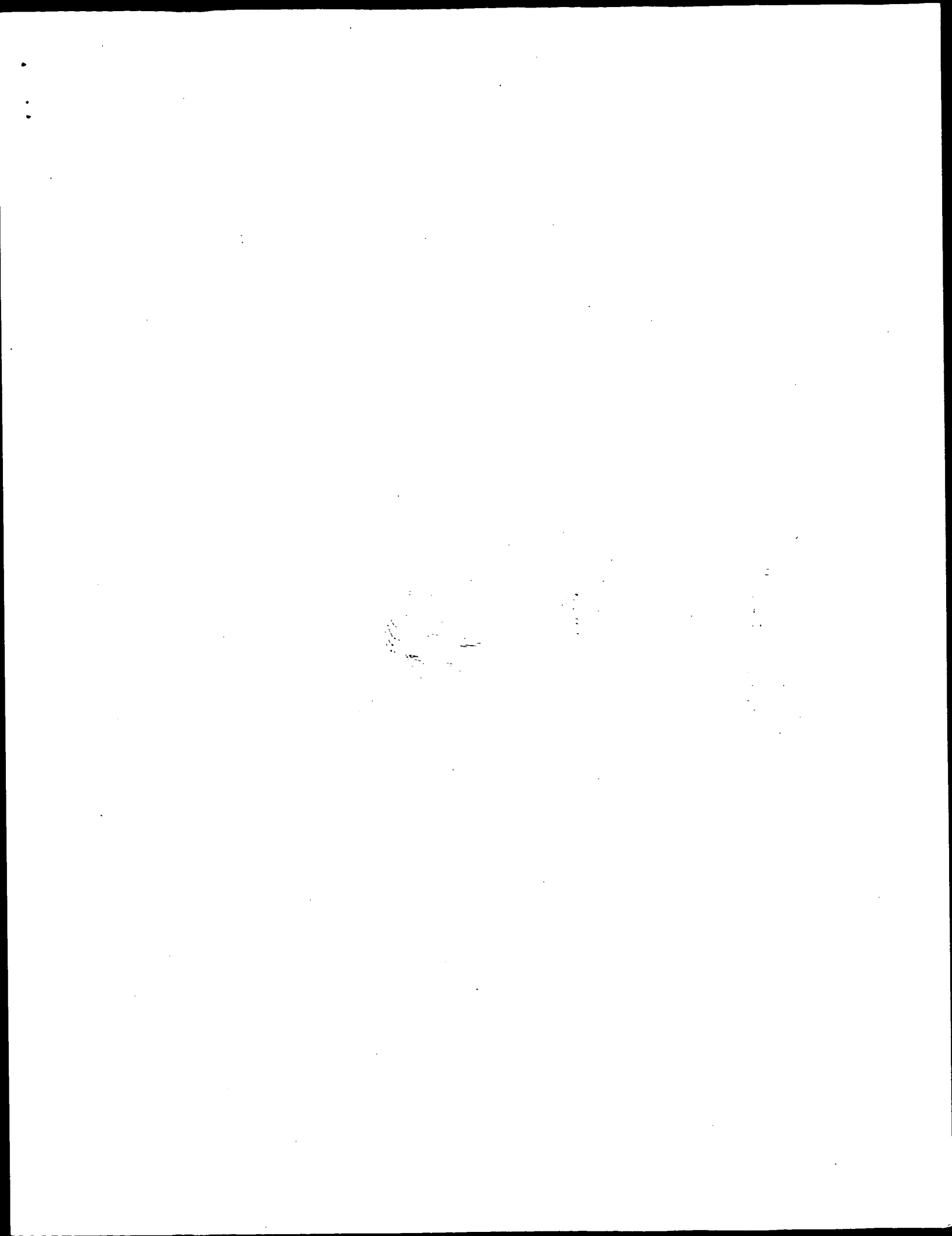
QY 4 EBAKTYLDFKFNHEADLFYQSSIASWNTNTTEENYONMANNAGDMSAFLEKQSTL 63
 DB 22 EBAKTYLDFKFNHEADLFYQSSIASWNTNTTEENYONMANNAGDMSAFLEKQSTL 63
 QY 64 MYPLQETIONLTVLQALQONSSVLSDESKRLNTLNTMSTYSGKVCNPDNPQ 123
 DB 82 MYPLQETIONLTVLQALQONSSVLSDESKRLNTLNTMSTYSGKVCNPDNPQ 123
 QY 82 MYPLQETIONLTVLQALQONSSVLSDESKRLNTLNTMSTYSGKVCNPDNPQ 123
 DB 82 MYPLQETIONLTVLQALQONSSVLSDESKRLNTLNTMSTYSGKVCNPDNPQ 123
 QY 124 -LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPIYEEYVLKNEARAHNEDY 182
 DB 142 -LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPIYEEYVLKNEARAHNEDY 182
 QY 142 -LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPIYEEYVLKNEARAHNEDY 182
 DB 142 -LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPIYEEYVLKNEARAHNEDY 182
 QY 183 DYMKGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 241
 DB 202 DYMKGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 241
 QY 202 DYMKGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 241
 DB 202 DYMKGDEVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHLAAVYRAKLMNAPPS-YTSP 241
 QY 242 GCLPAILGDMGCFWNTLSLVPPGOKPNDIVTAMDOAMDORIKREKFEVSV 301
 DB 252 GCLPAILGDMGCFWNTLSLVPPGOKPNDIVTAMDOAMDORIKREKFEVSV 301
 QY 252 GCLPAILGDMGCFWNTLSLVPPGOKPNDIVTAMDOAMDORIKREKFEVSV 301
 DB 252 GCLPAILGDMGCFWNTLSLVPPGOKPNDIVTAMDOAMDORIKREKFEVSV 301
 QY 302 LPMNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 DB 312 LPMNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 QY 312 LPMNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 DB 312 LPMNTQGFWSNLTDPGNVQAKVCHPTAMDLSGK-DPRILMCTRYMDPFLTAHENG 359
 QY 361 IYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 DB 372 IYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 QY 372 IYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 DB 372 IYDMAVYAOPLFLRNGANGFHEAVGELMSISATPKLKSIGLSPFOEDNETEIN 419
 QY 421 LKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 DB 431 LKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 QY 431 LKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 DB 431 LKQALTYGTLPTFYMELKRWMMVFKGEIPKQDMKKWEMKREIVGVVEPVHDETY 479
 QY 481 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 534
 DB 491 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 534
 QY 491 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 534
 DB 491 CDPASLFHVSNDYSFIRYTRITLXQFOFALQAAKHGSPHAKCDISNSTRAGQKLFEM 534

ID	AA	AA70013 standard; Protein; 660 AA.
DT	25-SEP-1995	(first entry)
DE	Tick carboxypeptidase.	
XX	Tick; antigen; carboxypeptidase; vaccine.	
XX	Boophilus microplus.	
XX	Key	Location/Qualifiers
FT	Peptide	30..53
FT	Peptide	/label= BM91 peptide
FT	Peptide	185..201
FT	Peptide	/label= peptide AAT91251
FT	Peptide	202..209
FT	Peptide	/label= AAT91141
FT	Peptide	225..253
FT	Peptide	/label= T9118
FT	Peptide	262..269
FT	Peptide	/label= Bm91 peptide
FT	Peptide	442..452
FT	Peptide	/label= T9129(a)
FT	Peptide	456..481
FT	Peptide	/label= T9129(b)
FT	Peptide	599..606
FT	Peptide	/label= T9109
FT	Peptide	1..29
FT	Domain	/label= signal
FT	Domain	639..655
FT	Domain	/label= C-terminal transmembrane domain
FT	Domain	/note= "putative"
XX	WO9504827-A.	
XX	16-FEB-1995.	
XX	10-AUG-1994;	94WO-AU00463.
XX	10-AUG-1993;	93AU-0000458.
XX	(BIOT-) BIOTECH AUSTRALIA PTY LTD.	
XX	(CSIR) COMMONWEALTH SCI & IND RES ORG.	
XX	Cobon GS, Kemp DH, Teliam RL, Willadsen P;	
XX	WP1; 1995-090905/12.	
XX	N-PSDB; AAQ82948.	
XX	New antigenic tick carboxypeptidase and corresp. DNA - are used	
XX	in vaccines for producing antibodies against ticks, insects and	
XX	nematodes.	
XX	Disclosure; Fig 6; 138pp; English.	
XX	Clone A5 was prep'd. from adult tick cDNA library. Clone 4U1 was	
XX	prep'd. from the larval stage of B. microplus (Calliope strain).	
XX	AAQ82948 is a hybrid of sequences from clone 4U1 (nt 1-966 & 1747-2047)	
XX	and A5 (nts 967-1746). The translation of the tick carboxypeptidase	
XX	cDNA sequence is shown in AA70013. All the native tick	
XX	carboxypeptidase sequences listed in Table 11 (see AA70014-R70023)	
XX	are found in the translation (see FT). The predicted AA sequence	
XX	agrees with the peptide sequence for all peptides with 2 exceptions.	
XX	These differences are Asp for Gln14 in peptide T9126, and Asn for	

```

CC Asp12 in peptide T9118. Tick carboxypeptidase has a mol. wt. 75172.
CC In addition to the features in FT, the carboxypeptidase AA sequence
CC has 8 potential N-linked glycosylation sites and a potential
CC glycosylphosphatidyl inositol anchor sequence similar to that found
CC in hmb6. It has significant homology with zinc dependent dipeptidyl
CC carboxypeptidases from mammals.
XX
SQ Sequence 660 AA;
Query Match 27.2%; Score 1054; DB 16; Length 660;
Best Local Similarity 37.7%; Pred. No. 2,5e-84;
Matches 224; Conservative 104; Mismatches 234; Indels 32; Gaps 12
QY 4 EEOAKFTLDFKFNHEADLFYQSSIASWNTNTTEFNQNMNN-AGDKMSAFILKQSTLA 62
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 49 EAMGVAFIIGLNDPYTTINNVDSSSSWDYASNTIDN-QMSNKVKSTESKNERQFGITA 107
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 63 QMPLDIEIQLTVKIQLOALQONGSSVISEDKSKRLTILNTMSTISYTGKVC---NPDN 119
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 108 KKFDMHNFNDKSLKLFREHATIGTALAIPDDKLENAITSLSKMAIYGSTKTVGKMDL 167
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 120 POECILIEPGLMEIANSIDYNERIMAWMSKSESGKQRLPYEEVYLKNNMAANNYE 179
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 168 F-----LEEDLIRNKKVEGNYTKLQITWLAIMNAVGPAIKQYIPIYIKISNNAASLDG 222
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 180 DYGDYWRGDEYVNGVDYDYSRGQILIEVHTEFEERIKPLYEHIAHYRAKLNNAPSYIS 239
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 223 NIKSAMLSDEYE-----NMTEIYDKLMEDLSPLYKKLHAHYMKLREIYGRLP 272
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 240 PIGCIIPAHILGMMGRFTWNLV-SLTYPGGKRPNDYVDAMQDQAMDQRIFFKEAEFFV 298
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 273 EDGTFPAHLIGMMMAQEWGTLPRILME--DKP-LDISKTAWEQWDAQKMFHADEFFT 329
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 299 SVGLERNMQGFWENSMITDPGNVQKAVCHPTAMDGKG-DPFIIMCTKVTMDPFLTAHE 357
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 330 SIGLNMNTSEFSKSIILTKPED-REIQCHASAMNNYNDDEPRIMKCTDPSVELRTVIHE 388
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 358 MGHIDYDAAVYAAPRLIRNGANGEGHEAVGELMSIAATPKHLSIGLSPDQDNTE 417
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 369 MGHIIYQVYKHLHVLLOGANGEGHEAVGDIILASVATKTHYGLSLIKP---TDKRYNA 445
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 418 INFLIKQALLTVGLDFTYMLTEKWMWYEFKGEIPDKQMMKRWMEKREIYGVPEVPYDE 477
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 446 VDLILMSALDKIAIAFLPFGYLLDKRMWTLFTGETPRDKNNEKFWERKITQGVSPYKRN 505
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 478 TYCDPASLFIHNSNDYSIFRTYTRITLYQGFQDALCOQAK---HSGPLHKCDISNSTEAG 534
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 506 SFEDGAGATYHALHVPYLRVFAALLDQRFENHLCYAKKADDEHHPFECDIYGEKNA 565
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 535 KLFNMLRKSGKEPWTIALENYVGAKKNNVRLYVBEPLFWLMDQNNSPVGM 588
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 566 VLKNGLSIGRSKPMPPVDLEIMAGTRQMSASSLKYIEPLKWLDBERIKNEVGM 619
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:53:54 ; Search time 22 Seconds

(without alignments)
962,932 Million cell updates/sec.

Title: US-09-978-385-2_COPY_19_738

Perfect score: 3869

Sequence: 1 STIEQAKTFIDKRNHAEAD.....DNSLEFLGIQPTLGPNNPP 720

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	805	4	US-08-989-299-2
2	1337	34.6	732	1	US-08-481-626-2
3	1337	34.6	732	4	US-08-989-299-4
4	1337	34.6	732	4	US-08-989-299-7
5	1334	34.5	732	4	US-08-989-299-5
6	1334	34.5	732	4	US-08-989-299-8
7	1310	33.9	737	4	US-08-989-299-6
8	1283	33.2	737	4	US-08-989-299-10
9	1283	33.2	737	4	US-08-989-299-11
10	1073	27.7	615	4	US-09-440-325A-1
11	989	25.6	907	4	US-08-989-299-12
12	638.5	16.5	907	4	US-08-989-299-15
13	120	3.1	990	2	US-08-630-916A-46
14	116	3.0	990	2	US-08-392-625-20
15	115	3.0	990	2	US-08-466-961A-20
16	113.5	2.9	665	3	US-08-844-059-2
17	113.5	2.9	665	3	US-09-431-202-2
18	104	2.7	834	1	US-08-491-357-2
19	104	2.7	834	1	US-08-968-633-2
20	104	2.7	834	1	US-09-196-466-2
21	104	2.7	834	1	US-08-968-633-2
22	104	2.7	834	1	US-08-968-633-2
23	102.5	2.6	631	4	US-08-448-489-17
24	102.5	2.6	631	4	US-08-704-711A-18
25	102.5	2.6	660	3	US-09-521-220-18
26	102.5	2.6	660	3	US-09-391-104-19
27	102	2.6	3169	4	US-09-453-702B-257

28	101.5	2.6	450	4	US-09-306-593-13	Sequence 13, Appl
29	101.5	2.6	560	4	US-09-134-001C-3153	Sequence 3153, Ap
30	99.5	2.6	930	4	US-08-953-040-2	Sequence 2, Appl
31	99	2.6	789	4	US-08-902-285-84	Sequence 84, Appl
32	97.5	2.5	570	3	US-08-747-221B-55	Sequence 55, Appl
33	97.5	2.5	570	4	US-09-005-051-55	Sequence 55, Appl
34	97.5	2.5	595	3	US-08-747-221B-31	Sequence 31, Appl
35	97.5	2.5	595	4	US-09-005-051-31	Sequence 31, Appl
36	97	2.5	977	3	US-08-335-844A-22	Sequence 22, Appl
37	96	2.5	789	4	US-09-002-285-82	Sequence 82, Appl
38	96	2.5	806	4	US-08-945-983-2	Sequence 2, Appl
39	96	2.5	3443	2	US-08-416-603-2	Sequence 2, Appl
40	94	2.4	452	1	US-08-317-880-2	Sequence 2, Appl
41	94	2.4	452	2	US-08-782-396-2	Sequence 2, Appl
42	93	2.4	647	2	US-08-844-056-2	Sequence 2, Appl
43	93	2.4	789	4	US-09-002-285-78	Sequence 78, Appl
44	93	2.4	789	4	US-09-002-285-94	Sequence 94, Appl
45	93	2.4	790	4	US-08-960-780-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-989-299-2
Sequence 2, Application US/08989299
Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Robinson, Susan L.
TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOGY
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: POLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 805 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein

Query Match 100.0%; Score 3869; DB 4; Length 805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 STIEQAKTFIDKRNHAEADLFYQSSLASWNTNTIENYQMMNAGDKWSAFLEKQST 60
19 STIEQAKTFIDKRNHAEADLFYQSSLASWNTNTIENYQMMNAGDKWSAFLEKQST 78
61 LAQMTPLQEIQNTLVKIQIQALQNGSSVLSDEKSKRLNTIINTMTSTIYSTRKVCNPDP 120

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/656,183
FILING DATE: 04-MAR-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 89-09062
FILING DATE: 05-JUL-1989
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 04958-0006-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-626-2

Query Match          34.6%; Score 1337; DB 1; Length 732;
Best Local Similarity 41.7%; Pred. No. 1,16-119;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9

QY      2 TTEQAKTFLDPEFNHEADLFYQSSLASMYNTNTITEE-----NVQNNNAGDKMSA 53
DB      70 TDAESKSVVEEDRISQVYWEYEAAMNNNTNTTETSKILLQKNQJANHT----- 123
QY      54 FKEQSTLAQMTPLQEIQLTVKQLQLQALQONGSSVISEDKSKLNTIITNTSTIGTK 113
DB      124 --LKYGQARKPEYDNLQMTTITKRIKKYQDLERAAIPQAQLEYYNILLDMETTSVAT 181
QY      114 VCPNDQNOQECILPEPLINEMANSIDYNERLMAWESRSEYQKQRLPEYEEYVLKNEMA 173
DB      182 VCHNG--SCLQLEPDLTVMAATSKRYEDLMAWGMCDKRGRIILQFYRYVELLIQAA 239
QY      174 RANHYEDYGDYWRGDEYVNGVDGYDSRGQILIEDYEHTFEETKPLYEHLHAYRAKLMNA 233
DB      240 RLNRYVAGSGWSRSMETPSLE-----QDLERLFOELQPLYLMLHAYRARALHRH 289
QY      234 Y-PEYISPICSLPAHLIGDMGGRFNTLSLTPVFGQKPNIDVDAMVQDAMQRIEKE 292
DB      290 YGAOHIDEGPIRAHLILGNMAQWTSNIITDLYPPPSAPSDTTEAMLKQGMTPRMFEKE 349
QY      293 AEKEFVSQGLPNNITQGFWENSMILDPGNVQKAVCHPTAMDLSG--DPEILMCTVITMDDF 351
DB      350 ADDEFTSISGLLPVPEFWMKSMLEKPTDREYVACHASANDYNGDFEIKQCTTVNLDEL 409
QY      352 LTAHHEGHILQYDAVYAAPFLRLNAGNAGEFVAQELISMAATPKHLISIGLLSPDQ 411
DB      410 VVAHHEGHIOYFQYQNDLPEALRBEGANGFHALGSDVALASVTPKHLISLNTLSSEG 465
QY      412 EDNETEINFLKQALITVGLPTPTYLEKRYKRVYEGEELPKQNMKKMKMKRELYAVE 471
DB      470 SD-EHDIIDILKMAALDKIAIPSYLVDQRMVWEFGSTIKENYQEMWLSRLKYGSLCP 528
QY      472 PVPHDETCDASLPEFHSNDYSFIRYRTYRLVYQFOFQALCQAAKHGEGFLHKCDISNTE 533
DB      528 PVPRKQGFDEGAFHFHPSSVPYIRYVPSIIIOFQHEALCQAAGHTGFLHKCDIYOSKE 586
QY      533 AGQILNMLRLGKSEPTLALENVYGAKKNNVPLRYLPEELPYLTKQDNK--NSFYGM- 580
DB      589 AGQRILATLMKIGFSRHPWEAMQILITGPNNMSASAMSLYKRLPDLWLTENLHGEKLGWP 644
QY      589 STDWSPYADS 599
DB      649 QYTWIPNSKRS 659

```


RESULT 3
US-08-989-299-4
Sequence 4, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-7000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-4

Query Match
Best Local Similarity 34.6%; Score 1337; DB 4; Length 732;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

2 TIEQAKTFLDKFNEHAEADLFYQSSLASMYNTNITEE-----NVQNNNAGDKWSA 53
70 TDEAASKFEVEYEDRTSOVYMNETAENAMNNTNITETSKILLQKNQIANHT-----123
54 FLKEQSTLAQMTPLQEIQNTLVKIQLOALQONGSSVLSDESKRLNTILNTMSTIYSTGK 113
124 --LKYGTQARKFDVNOJLQNTTIKRIKKVODLERALPAGOLEEYKNKILLDMETTVSVAT 181
114 VCPNDNPOECILLEPGLNEIMANSIDYNERLAMESRSEVKGQRLPYEEYVVLKKNEMA 173
182 VCHPNC--SCQLEPDLNVAATSRKYEEDLIMAWEGWRDAGRAIILQFYKYEELINQAA 239
174 RANHEDYGDYWRGDEYVNGDYDSHGLIEVEHTFEERIKPLHYHLAAVYRAKIMANA 233
240 RLNGYVADAGDSMRSMYEPSPLE-----QLEKLFQLEQLOPLVNLHAYVRALHRE 289
234 Y-PSYISPIGLPRAHLIGDMGRFWMLYSLTVPGQKPNIDYVDAVDAQDAQRIKFE 292
290 YGAQIINIEGPIPAHLIGNMMAQWMSNIYDLVVPFSAFMDTTEALMLKQGWPRRMEKE 349
293 AEKEFVSGLFPMTOGFENSMILTDPGNVOKAVCHPAMDLGK--DFRIIMCTKYVTMDF 351
350 ADDEFTSLGLLPVPEREWMKSMLEKPTDGEVYVCHASAMDFYNGKDFRIKQCTTVMLIDL 409
352 LTAHHEGHQIYDMAYIAQAPLLRNGANEGFHEAVEIMSLAATPKHLSIGLISPDQ 411
410 VVAHHEGHQIYDMAYIAQAPLLRNGANEGFHEAVEIMSLAATPKHLSIGLISPDQ 469

RESULT 4
US-08-989-299-7
Sequence 7, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-7000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1306 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-7

Query Match
Best Local Similarity 34.6%; Score 1337; DB 4; Length 1306;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

2 TIEQAKTFLDKFNEHAEADLFYQSSLASMYNTNITEE-----NVQNNNAGDKWSA 53
644 TDEAASKFEVEYEDRTSOVYMNETAENAMNNTNITETSKILLQKNQIANHT-----697
54 FLKEQSTLAQMTPLQEIQNTLVKIQLOALQONGSSVLSDESKRLNTILNTMSTIYSTGK 113
698 --LKYGTQARKFDVNOJLQNTTIKRIKKVODLERALPAGOLEEYKNKILLDMETTVSVAT 755
114 VCPNDNPOECILLEPGLNEIMANSIDYNERLAMESRSEVKGQRLPYEEYVVLKKNEMA 173

412 EDNETEINFLKQALITVLTPTMYLMEKRWMTVEKGEIPDKDMKKMKREIYGYVE 471
470 SD-EHDINFLMKMALDKIAPIFSYIVDQMRKRPDSITREKYNDEWMSLRUKYGLCP 528
472 PVPHEETCDPASLFHVSNDYSFTRITRYLQFOFQALCOAKHEGPHLKCDISNSTE 531
529 PVPRTQGFDPGAFHFISSVPYIRYVSFIQFOFHEALCOAGHGTGLKCDIQSKE 588
532 AGOKLFNNLRIGKSEPTLLENVYGAKNMVRPLNYPRLPTWTKDNK--NSFVGA- 588
589 AGORLATAMKLGSRPPEMAMQLTGQPNMSASAMLSYERPLDMLTEHELGERLGP 648
589 STDMSPYADQS 599
649 QYNWTPNSARS 659

```

DB 756 VCHNG--SCQLEPDLTNMATSRYEDLLMAMEGMRKAGRAIIOFPKVEYLINQAA 813
OY 174 RANIEDYGDYRDYEVNGVDYDSRGQLEDVEHTEFEIKPLXEHLHAYRAKLMA 233
DB 814 RLNGYVAGDSWRSMYETPSLE-----QDLERLQLOPLNLHAYRRALNH 863
OY 224 Y-PSYISPIGCLPAHLIDGMGRFTNLSLTVPEGQKNIDVTADMDAORIEFE 292
DB 864 YGACHINIEGIPRPHILGNNMAQWTSNIDLVYFPSPASMDYTEAMLKQWTPRMEKE 923
OY 293 AEKEFVSGLDPLNMTQGWENSMLTDPGNQKACHPTAMDLSGK-DERILMCTKYTMDPF 351
DB 924 ADDEFTSLGILPVPEEFWKNKSMLEKPTDGREVYCHASAMPYNGKDPRIKQCTVMEBL 983
OY 352 LTAHEMGHIQDMAVAQAOPFLRNGANEGHEAVGIMSLSAATPKHLKSTIGLSPDFQ 411
DB 984 VVAHEMGHIQYFMQYKDLPLALREGANPGHEALIGVYLAISYTPKHLSLMLSEGG 1043
OY 412 EDNETEINFLKQALITVGLPFTYMLEKRMWVFKEIIPKDOMMKWMEKREIYGYE 471
DB 1044 SD-EHDINFLMKMALDKIAFIPEISYLDQMRWVFDGSIKENYQMSLRKLYQGLCP 1102
OY 472 PVPHEBTCDPASLEPHVSNDSFIRYTRTLYOFOFOBALCOAKHSGPLKCDISNTE 531
DB 1103 PVPRTQGDPEAKHIDSVPYIRYFVSFIIOFPHALCOAGHTGPHKCDIYOSKE 1162
OY 532 AQOKLENNLRKSEPTLALENVGAKNNVRPLNTEPFLTWLKDOK--NSFVGW- 568
DB 1163 AQORLATAMKLGFSRPEAMOLITGQNMASASMLSYFKPLDMLTLENLHGEKLGMP 1222
OY 568 STDMSPYADOS 599
DB 1223 QYNMTPNSARS 1233

RESULT 5
US-08-989-299-5
; Sequence 5, Application US/08989299
; Patent No. 6194536
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-5
Query Match 34.5%; Score 1334; DB 4; Length 732;
Best Local Similarity 42.6%; Pred. No. 2,2e-119; Indels 18; Gaps 7;
Matches 255; Conservative 112; Mismatches 213;

OY 2 TIEQATFLDKFHEAEFLYQSLSANYNINTEEVQNMNNAQKSAFLKEOSTL 61
DB 69 TDEAKRPFEEYDRTAQVILNEYAEANMOYNINITEGSKILLESTVSHHTLYGR 128
OY 62 AQMTPLQIONLTVKLOLAQOQSSVLSDESKRLNTLTMTSTTSTGYVCPDNQ 121
DB 129 AKTDVSNFQSSIKRIKTLQNDRAVLPKLEEVNOILLDMETYSLSINCYTNG-- 186
OY 122 ECLLEPGLNFIANSIDYNERLMAWSMSSEYQKQRLPYEYVYLKMMARANYEDY 181
DB 187 TCMPLRPLTNMMASTRKYBELLMAMKSMKDKGRALIPFPKYVEFSNKLAKNGYTDA 246
OY 182 GDYWRGDYEVNGVDYDSRGQLEDVEHTEFEIKPLXEHLHAYRAKLMAVPS-YISP 240
DB 247 GDSMSLYESDNL-----QDLKLYQLOPLYLNLHAYVRSLSRHYSSEYINL 296
OY 241 IGCLEPAHLIDGMGRFTNLSLTVPEGQKNIDVTADMDAORIEFEKEKEFVS 300
DB 297 DGPIDPAHLIDGMMAQWTSNIDLVAPPASANIDATAMIKQGTPTRIKTEADNFTSL 356
OY 301 GLPNMTQGWENSMLTDPGNQKACHPTAMDLSGK-DERILMCTKYTMDPFLAHENG 359
DB 357 GLPVPEEFWKNKSMLEKPTDGREVYCHASAMPYNGKDPRIKQCTVMEBLVIAHENG 416
OY 360 HIQDMAVAQAOPFLRNGANEGHEAVGIMSLSAATPKHLKSTIGLSPDFQEDNEIN 419
DB 417 HIQYFMQYKDLPLALREGANPGHEALIGVYLAISYTPKHLSLMLSEGG 475
OY 420 FILKQALITVGLPFTYMLEKRMWVFKEIIPKDOMMKWMEKREIYGYEVPHEBT 479
DB 476 FLKMAKLDKIAFIPEISYLDQMRWVFDGSIKENYQMSLRKLYQGLCPVPBSQGD 535
OY 480 CDPASLEPHVSNDSFIRYTRTLYOFOFOBALCOAKHSGPLKCDISNTEAQOKLFPM 539
DB 536 FDPGSKFHVPAWVYRYVFSFIIOFPHALCRAAGHTGPHKCDIYOSKEAKLADA 595
OY 540 LRLKSEPTLALENVGAKNNVRPLNTEPFLTWLKDOK--NSFVGW-STDMS 594
DB 596 MKLGSKPWEAKLITGQNMASASAMNTEKPLTEMLVTEHRRRGTGLGPEYINAP 653

RESULT 6
US-08-989-299-8
; Sequence 8, Application US/08989299
; Patent No. 6194536
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-8

Query Match 34.5%; Score 1334; DB 4; Length 1312;
Best Local Similarity 42.0%; Pred. No. 5,8e-119;

Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 2 TIEOAKTFLEKNEHAEEDLFYQSSLASWNTNTITEENVQNNNAAGKWSAFLEKQSTL 61
DB 649 TDEAKADRYEEDRTAQLVLENYEAMNQYNTNITIEGSKILKESTEVSHHTLKGTGR 708
QY 62 AOMYPLQELIONTVLQLOALOONGSVLSSEDSKRLNTILNTMSTIYSGKVCNPDNQ 121
DB 709 AKTFDVSNFQSSIKRIKIKQLNDRAVLPKELEBYNQILDMEITYSLNICTNG-- 766
QY 122 ECLLEPGLEIMANSIDYNERLWAMESRSEVQOLRPLYEEYVVKEMARANHEDY 181
DB 767 TCMPEPLDTLNMMAQTSRKEELLMAMKSWDRKVRGRLIPFPKYVEFSKIKLNGYDA 826
QY 182 GDYMRGDEYVNGVDYDSRGQLIEDVEHTFEIKPLYLEHLAVYRAKLMNAPS-YISP 240
DB 827 GDSWRSYESSDLE-----QDEKLYQELQPLYLNLHAYVRSLSHRHYGSEYINL 876
QY 241 ICGPLAHLLGDMGRFTNLYSLVPPGQPNIDVTDAWDQAMDORIKREAEKFFVS 300
DB 877 DGPPIAHLLGDMMAQTSWNTYDLVAPPSAPNIDATEAMIKQGWPRIRFEKADNFTSL 936
QY 301 GLPMNTOGFWSNMLTDPGNVOKAVCHPTAMDLGK-DPRILMCTKYTMDFLTAHHEG 359
DB 937 GLPVPPEFWNKSMLKEKPTDGRVYCHPSAMDYNGKDFRKQCTSVNMEDLVIAHHEG 996
QY 360 HIQYDAMAAOPLFLRNGANEGFHEAVGEIMSLSATPKHLKSGILSPDQEDNETEIN 419
DB 997 HIQYFMQYKDLVTFREGANPGFHEALIGDIMALSVTRKHLTSLMLSTE-GSGYEYDIN 1055
QY 420 FLIKQALTYGTLPTTYMLEKRWAVFKEIPKQDMKRWEMKREIYGVVPEVPHDET 479
DB 1056 FLMKMALDKLAIFFSYLIDQWRHVFDSITKENYDQEMWLSRLKYQGLCPVPRSGD 1115
QY 480 CDPALEFVNSDYSPRIYRTYTRTLVQFOFQALCOAAKHGRLKCDISNTEAGOKLFNM 539
DB 1116 FDPGSEFHVANVPYRIFVSFIIDQFHEALCRAGHGTPLAKCDIYQSKAGKLADA 1175
QY 540 LRIKSEPTLLENVYGAKNNAVRPLLYFEPLFTWLKDNK--NSFVGM-STWSP 594
DB 1176 MKLGSKPWPPEAMKILITQPPNNSASAMNVEFKLEMLVTENRRRGEITLGNPEYIMAP 1233

RESULT 7
US-08-989-299-9
Sequence 9, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Acton, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOGENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,299
FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-9

Query Match 33.9%; Score 1310; DB 4; Length 1313;
Best Local Similarity 42.0%; Pred. No. 1,2e-116;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

QY 2 TIEOAKTFLEKNEHAEEDLFYQSSLASWNTNTITEENVQNNNAAGKWSAFLEKQSTL 61
DB 650 TDEAKANFVEYEDRTAKVLWNEYEAMNHNTNITIEGSKILQKKEYSNHTLKGTW 709
QY 62 AOMYPLQELIONTVLQLOALOONGSVLSSEDSKRLNTILNTMSTIYSGKVCNPDNQ 121
DB 710 AKTFDVSNFQSSIKRIKIKQVNDRAVLPNELEBYNQILDMEITYSLNICTNG-- 767
QY 122 ECLLEPGLEIMANSIDYNERLWAMESRSEVQOLRPLYEEYVVKEMARANHEDY 181
DB 768 TCMPEPLDTLNMMAQTSRKEELLMAMKSWDRKVRGRLIPFPKYVEFSKIKLNGYDA 827
QY 182 GDYMRGDEYVNGVDYDSRGQLIEDVEHTFEIKPLYLEHLAVYRAKLMNAPS-YISP 240
DB 828 GDSWRSYESSDLE-----QDEKLYQELQPLYLNLHAYVRSLSHRHYGSEYINL 877
QY 241 ICGPLAHLLGDMGRFTNLYSLVPPGQPNIDVTDAWDQAMDORIKREAEKFFVS 300
DB 878 DGPPIAHLLGDMMAQTSWNTYDLVAPPSAPNIDATEAMIKQGWPRIRFEKADNFTSL 937
QY 301 GLPMNTOGFWSNMLTDPGNVOKAVCHPTAMDLGK-DPRILMCTKYTMDFLTAHHEG 359
DB 938 GLPVPPEFWNKSMLKEKPTDGRVYCHPSAMDYNGKDFRKQCTSVNMEDLVIAHHEG 997
QY 360 HIQYDAMAAOPLFLRNGANEGFHEAVGEIMSLSATPKHLKSGILSPDQEDNETEIN 419
DB 998 HIQYFMQYKDLVTFREGANPGFHEALIGDIMALSVTRKHLTSLMLSTE-GSGYEYDIN 1056
QY 420 FLIKQALTYGTLPTTYMLEKRWAVFKEIPKQDMKRWEMKREIYGVVPEVPHDET 479
DB 1057 FLMKMALDKLAIFFSYLIDQWRHVFDSITKENYDQEMWLSRLKYQGLCPVPRSGD 1116
QY 480 CDPALEFVNSDYSPRIYRTYTRTLVQFOFQALCOAAKHGRLKCDISNTEAGOKLFNM 539
DB 1117 FDPGSEFHVANVPYRIFVSFIIDQFHEALCRAGHGTPLAKCDIYQSKAGKLADA 1176

QY 540 LRLKSEPTLALENVGAKNNVPLNTFEEPLTWLKDQNK--NSFVGM-STDNSP 594
 DB 1177 MKLGSKQWPEAKMTITGOPNMSASALMNYFKPLTEWLVENRRHGSLGMPXYTWP 1234

RESULT 8

US-08-989-299-6

Sequence 6, Application US/08989299

Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG

TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESS: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,299

FILING DATE: 11-DEC-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Arnold E., Beth

REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MIA-025.01

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 737 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-989-299-6

Query Match

Best Local Similarity 40.8%; Pred. No. 1.8e-114;

Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

QY 2 TIEBQAKTFLDKFNHAEADLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 51

DB 75 TDEAASRVEVEYDRSFOAVNMEYEAANNNTNTITEEASKILLQKNQJANHLLTYGNW 134

QY 52 SAFLKEOSTLAQMTPLQEIOMLVKLOLQONGSSVSEDSKRLNTIANTSTYST 111

DB 135 -----ARRDVSNFQNAATSKRIKKVODLQRAVLDPKKELEEVNQLIDMETIYSV 184

QY 112 GKVCNPDPQECILLEGELINMANSIDYNERLWAMESRSEVQKORLPYEEVVLKNE 171

DB 185 ANVCVDS--SCLDLEPDLNLMATSKRYDELWMTSMKQGRALLPFPKIVETPNK 242

QY 172 MARANHEDYDYNRGDYEVNGVDYDSRGLIEDVEHTFEELKPYEHLHAYVRAKIM 231

DB 243 AARLNGYDADDSRMSKETPTLE-----QDLERFQELQPLIXINMHAVALH 292

QY 232 NAY-PSYISPTGLCPAHLIGDMGRFMTNLYSLTFPGQKPNIDYDAMVQDAQRIIF 290

DB 293 RHYGAQHINTEGPIPAHLIGNMMAQOTSNITDVAAPPASATDMATEAMIKOGWTPRRMF 352

QY 291 KEAEKFEYSGLPMNTOGFEMNSMLTDPGNQKAVCHPLAMDQK--DERLIMQTKYTM 349

DB 648 TDEAASRVEVEYDRSFOAVNMEYEAANNNTNTITEEASKILLQKNQJANHLLTYGNW 707

DB 353 EEADEFFISLGLPVPPEEFWNKSMLEKPTDREYVCHASAMDFYNGKDFRIKQCTYVME 412
 QY 350 DELTAHENGHIQYDAAVAAOPFLRNGANEGFHEAVGEINSLSATPKHLKSLGLSPD 409
 DB 413 DLVYVHENGHIQYFQYKDLFVALREGANPGFHAIGDVALSLSTYKHLHSLNLSSE 472
 QY 410 FQEDNETEINFLKQALITVGLPFTYMLEKRWNVFGEIPLKQDMKRWEMKREIVGV 469
 DB 473 -GGGYHNDINFLMKMLDKIAFIPIPSYIVDEMRHVFDSITKENYNOEMWSLRLKYQGL 531
 QY 470 VEPHDEYCDPASLFHVSNDYSPFRYTRTLVQFOEALCOAKHKGPLHCDJNS 529
 DB 532 CPAPRSQGFDPGAKFHLPSSVPYIRFVSFTIOFHEALCKAAGHGPLHTCDIYQS 591
 QY 530 TEAGOKLFMLRLGKSEPTLALENVGAKNNVPLNTFEEPLTWLKDQNK--KNSFVG 587
 DB 592 KEAGKRLADAMKLGYSKPPWPAKMTITGOPNMSASAMNYPPLMDWLTLENGRGKLG 651
 QY 588 W-STDNSPYADOS 599
 DB 652 WQYTWTPNSARS 664

RESULT 9

US-08-989-299-10

Sequence 10, Application US/08989299

Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG

TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESS: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,299

FILING DATE: 11-DEC-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Arnold E., Beth

REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MIA-025.01

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1310 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-989-299-10

Query Match

Best Local Similarity 40.8%; Pred. No. 4.6e-114;

Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

QY 2 TIEBQAKTFLDKFNHAEADLFYQSSLASWNTNTITEE-----NVQNMNN--AGDKW 51

DB 648 TDEAASRVEVEYDRSFOAVNMEYEAANNNTNTITEEASKILLQKNQJANHLLTYGNW 707

```

OY 52 SAFLEOSTLAOMPILOEIONLVKLOALQONGSSVLSHDSKRLNTLNTMSTYST 111
DB 708 -----ARRVDNSFONATSKRIKKVODQRAVLVPEKLEENQILDMETISV 757
OY 112 GKVCNPNPOECILLEGLNEMANSIDYNERLWAMESWSEVQKOLRPLYEEVYLKNE 171
DB 758 ANVCNRVGS--SCLOLEDPLTLMATSKRYDELLMWTSMCKGRALLPFPKRYEPTNK 815
OY 172 MARANHEDYGDYRGDYEVNGDYDSRGOLIEVHTFEELKPLYEHLHAYRAKMA 231
DB 816 AARLNGYVDAGDSWRSMTETPLE-----QDLERLFOELQPIYLMHAYGALH 865
OY 232 NAY-PSYISPIGCLPALLDGMGRFTNTLSLTPFGQRPNTDVAWDQAMDQRIF 290
DB 866 RHIGAOHINLEGPRLALGNMMAQWTSNTIDYVAPPSASTMDATEAMTKQGTPEPRMF 925
OY 291 KEAEKFVSYGLPMTGFWENSMLTDPGVOKAVCHPTAMDGKG--DEFLMCTVTMD 349
DB 926 EADKPFISGLLPVPEFNNKSMLEKPTDGREVYCHASAMDPYNGKDFIKCCTVYVME 985
OY 350 DELTAHEMGHIYDAAVYAAOPFLLRNGANEHFEAVGELTMSLAATPKHLKSIGLSPD 409
DB 986 DLVYVHEMGHIYFMOKDLPVALREGANPGEHAIGDALVLSVSTPKHLHSINLSS 1045
OY 410 FOEDNETEINFLKQALTIYGLPFTYMLEKRMWYFKGPIPDQMKKWEKKREIVGY 469
DB 1046 -GGGEHNDINFLKMAKLDIAETPESTYLDWMRWVDSITKENTNOEWMSLRKYGL 1104
OY 470 VEREPHDEYCDPASLEHVSNDYSFTRYTRILYOFQFQALCOAKHEGPHKDISNS 529
DB 1105 CPAPRSQGDGDFGAEFLIPSSVPIRYEVSFIIOGFHEALCKAAGHGPCHTCIYOS 1164
OY 530 TEAGOKLEMLRKGSEPTWLALENVVGAKNNVRLNTEPELFTWLKDN--KNSFWG 587
DB 1165 KEKGKRLADMKGYSKPWEAKVITGPNMSASAMNYFKPLMDLLENGRHEKIG 1224
OY 588 W-STWSPYADQS 599
DB 1225 WFOYTWTPNSARS 1237

RESULT 10
US-08-989-299-11
; Sequence 11, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MTA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000

```

```

; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-11

Query Match 27.7%; Score 1073; DB 4; Length 615;
Best Local Similarity 36.5%; Pred. No. 2,2e-94;
Matches 218; Conservative 120; Mismatches 238; Indels 22; Gaps 9;

OY 4 EQGAKTFIDKFNHEADLEFYQSSLSWNTNTNITEBNVONNNAQDKWSAFLEQSTLAQ 63
DB 22 ETOAKEEYLENLKELAKRTNVETEAMAGNSITENEKKEKKEISAEELAKFEKVASDPT 81
OY 64 MYPLDEIONLVKLOALQONGSSVLSHDSKRLNTLNTMSTYSTGKYONPNPOEC 123
DB 82 KFGWRSYQSEDLKROKALTKIGYALPEDDYAELLDTLSMESNFAKVCYCDYKDSKC 141
OY 124 -LLEPEGLNEMANSIDYNERLWAMESWSEVQKOLRPLYEEVYLKNNMARANHEDYG 182
DB 142 DLALPELEBEVLSKSRDHEELAYWREFYDKAGTAVRSOFERYVELNTRKAKLNFTGA 201
OY 183 DYWRGDEVNGVDGYDSRGOLIEVHTFEELKPLYEHLHAYRAKLMNY-PSYISPI 241
DB 202 EAMLEDEY-----DITFEQDLEDI---FADIRPLYOOIHGYAFRLRKHYGAVVSET 251
OY 242 GCLPAILLGDGMGRFTNTLSLTPFGQRPNTDVAWDQAMDQRIFKAEKFEVSVG 301
DB 252 GPIPMHLGNMMAQWSEIADIVSPFKPLVDYSAEMEKQAYPLKFGQDDFFTSAN 311
OY 302 LPMNQGWENSMLTDPGVOKAVCHPTAMDGK--GDFRIIMCTKYTMDELTAHEMGH 360
DB 312 LTKLPQDWFWDSTIEKPPDGRDLVCHASAMDPYLLDVRIRKQCRVYTDOLFTVHHELG 371
OY 361 IOYDMAVAAOPFLLRNGANEHFEAVGELTMSLAATPKHLKSIGLSPDEDETEINF 420
DB 372 IOYFLQYOHOPRYVTGTANPGFHEHVGVLSTVSTPHLEKIGLKL-DYVRDEEARINQ 430
OY 421 LKQALTIYGLPFTYMLEKRMWYFKGPIPDQMKKWEKKREIVGVVPHDETYC 480
DB 431 LFLTALDLYVLEFATMDKRYMSLFRGEVDKAMNNCAFWLRDREYSIEPPVRSKDF 490
OY 481 DPASLEHVSNDYSFTRYTRILYOFQFQALC-QAAKH-----BGPLHKDISNTEAGQ 534
DB 491 DAPAKTHLSADVEYRKYVVSFTIOFQFKSACIAGQYDPDNLPLDNCIDYGSARAGA 550
OY 535 KLENNLRKGSEPTWLALENVVGAKNNVRLNTEPELFTWLKDN--KNSFWGMS 590
DB 551 AFHNNLSGASKPMPMDLLEAFNGERINSGKAIATFEPLRWLEAENIKNNVHIGWTT 608

RESULT 11
US-09-440-325A-1
; Sequence 1, Application US/09440325A
; Patent No. 6280994
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Zaci: A Human Metalloenzyme
; FILE REFERENCE: 98-79
; CURRENT APPLICATION NUMBER: US/09/440,325A
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT

```

LOCATION: (1)...(694)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-440-325A-1

Query Match 25.6%; Score 989; DB 4; Length 694;
Best Local Similarity 37.9%; Pred. No. 3.2e-86;
Matches 221; Conservative 104; Mismatches 228; Indels 30; Gaps 15;

QY 4 EEOAKTEFLDKFNHAEADLFYQSLSASWNTNTEENVOQNNNAGDKMSAFLEQSTLAQ 63
DB 53 ETEETKTEFLQFYDGTGVNLKMEKMTNVTNITKNOPEMKMDER-SQFPIITGTQAH 111
QY 64 MYPLAEIONLTVKLOLALQNGSSVLSDEKSKRLNTLNTSTITSTGKYCNPDPQEC 123
DB 112 LKRVYQFKPDVNGMLSKLQNDIKALSKDELREYNELLAXLEMTYSMAQVLNCGP-C 169
QY 124 LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLEEYVYLKNEMARAHN-YEYUG 182
DB 170 LSLSESL-EYMATSRDEKELLMWQMDVAGRQCTTTEHYVELSKRAQVLNGLYKXKDG 228
QY 183 DVMRGDYEVNGVDYDYSRGQLEDEYHTEFEIKRPLEYELHAYRAKLNNAY-PSYISPI 241
DB 229 ALMHSYTESDTLE-----QDEERLFOELRPLYLNPHTYVRALHRRHGPEDIDR 278
QY 242 GCLPAHLIGD-MMGREWTNLSLTYPEGKPNIDVTDAVDAQDAQRI-FKEAEKEFVS 299
DB 279 GPPLAHLIGENTLAQSWNILDVLPFLKIPEDVTYKIMKVOHWKPEKIMLEAEETFEY 338
QY 300 VG--LERNMTOGFENSMILDPGVNOKAVCHPRAMDLGK-GDRILLCTVTMDPFLTAH 356
DB 339 LGIATLPAPSEFKKMLMKPDTGREVCHISAMNYODDERIKKCAVTTEDPLSTH 398
QY 357 EMGHIOYDAAVYAAQPLLRNGANEHGEHVAEGLMSAATPKHLKSIGLSLDPQEDNET 416
DB 399 EMGHIOYDAAVYAAQPLLRNGANEHGEHVAEGLMSAATPKHLKSIGLSLDPQEDNET 453
QY 417 EINFELKQALITVGLPFTYMLEKMTNVTNITKNOPEMKMDER-SQFPIITGTQAH 475
DB 454 EVNFMHIALEKIAFTPFGLTMDLFRWKVFDGTIMWDIYQEWNNLRRLKYQGLCPALH 513
QY 476 DENYCPASLFHYSNYSFI-RYVTRTYLQFQFQALCOAAKHEGLHCDISNSTEAGQ 534
DB 514 SEEDPFGAKFHNSAGPYTRKRFELSYLVLOFQFHELCKRASGMGLHCDIDYNSIAK 573
QY 535 KLFNMLRLKSEFWTLALENVGAKNNVRLINTEPELFTWL 577
DB 574 LL--ALKIGSSKPPPEVLKMLTGESEYSTNVTYFKPLLTWL 614

RESULT 12
US-08-989-299-12
Sequence 12, Application US/08989299
Patent No. 6194556
GENERAL INFORMATION:
APPLICANT: Action, Susan L.
APPLICANT: Robinson, Keith E.
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989, 299

FILING DATE: 11-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Arnold E., Beth
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MIA-025.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-989-299-12

Query Match 16.5%; Score 638.5; DB 4; Length 907;
Best Local Similarity 27.0%; Pred. No. 2.6e-52;
Matches 167; Conservative 132; Mismatches 280; Indels 39; Gaps 18;

QY 4 EEOAKTEFLDKFNHAEADLFYQSLSASWNTNTEENVOQNNNAGDKMSAFLEQSTLAQ 63
DB 178 EELKRLMAGYEAELALIVLRVALSGRYPNDASPLKALDEANVLTFMVSTSMQAK 237
QY 64 MYPLAEIONLTVKLOLALQNGSSVLSDEKSKRLNTLNTSTITSTGKYCNPDPQEC 123
DB 238 QEDMASVDEKVMKLGVSFEFGMSALAPSRFADYSQAQALNDSKSTICDKDVPFC 297
QY 124 LLEPGLNEIMANSIDYNERLWAMESRSEVGKOLRPLEEYVYLKNEMARAHNEDYGD 183
DB 298 ALQKIDMOSITRNEKQASRLQHLWVSYTAIAKS-KFSYNNITISNKGAKNGFANGA 356
QY 184 YWRGDYEVNG-VDGYDYSRGQLEDEYHTEFEIKRPLEYELHAYRAKLNNAY-PSYISPI 240
DB 357 MWRSAFDMSSKYHKAFF--DLNKQIDKISTIQPFQOLLAHAYRRDLAGIYSNPVLSK 413
QY 241 IGCPLAHLIGDMKGFNTNLSLTYPEGKPNIDVTDAVDAQDAQRI-FKEAEKEF 296
DB 414 DGPPLAHLIGSLDGDMSAHYEQKPFEEES--ETPEAMLSAFTONTTKKMTVTAIRY 471
QY 297 FVSVGLPMMTOGFENSMILDPGVNOKAVCHP-TAMDL-GKDRILLCTVTMDPFLTAH 354
DB 472 FKSAGFPHLPKSYWTSIFARVWS-KMICHPAALDRAVNDPRVYKACAGLPDEQA 530
QY 355 HHEMGHIOYDAAVYAAQPLLRNGANEHGEHVAEGLMSAATPKHLKSIGLSLDPQEDNET 414
DB 531 HSLVQTYQYLYLKQDSLLFREQASPVITDAIANAFHLSNPHYLSQKLVSEHDIK 590
QY 415 ETE-ENFLKQALITVGLPFTYMLEKMTNVTNITKNOPEMKMDER-SQFPIITGTQAH 475
DB 591 DSVIINKLKRESLEFETKPIITADNNRYELFDGTIVKNNLNDRWMEIRNKYGVSPQ 650
QY 474 PHDEYCPASLFH--VSNDYSFIRYRTL---YQFQDALCOAA--KHGGLPLK 524
DB 651 PYNTSMLD--ALIHNSVGSVHS---PARRTLISYALKQIICALCORLFWLS---C 701
QY 525 DISNSTEAGQKLFNMLRLKSEFWTLALENVGAKNNVRLINTEPELFTWLKQDNK-- 582
DB 702 TLSEDTT-EKLRTEKLGSSITWLKALEMISGKGLDQAPLEYEPLINLNRNTNEID 759
QY 583 NSFVGSWTDSPYADSI 600
DB 760 QVYVGWGDGEGTPEYVEI 777

RESULT 13
US-08-645-193B-15
Sequence 15, Application US/08645193B
Patent No. 5962253
GENERAL INFORMATION:
APPLICANT: Kupke, Thomas

APPLICANT: Gutz, Friedrich
 APPLICANT: Kempter, Christoph
 APPLICANT: Jung, Gunther
 TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
 TITLE OF INVENTION: Catalyzed by Flavoprotein EpiD
 NUMBER OF SEQUENCES: 70
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/645,193B
 FILING DATE: 13-MAY-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652,1540000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2540
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 990 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 US-08-645-193B-15

Query Match 3.1%; Score 120; DB 2; Length 990;
 Best Local Similarity 18.2%; Pred. No. 0.023;
 Matches 150; Conservative 116; Mismatches 246; Indels 310; Gaps 40;

QY 2 TIEBOATFLDKFNH-----ADL--FIOSSIASWNYMTTEENVOMNNAADKW 51
 DB 343 TINIHEFFDKTGFQOLVNLKQSLDINGFGPKKDYSEFNNI----- 387
 QY 52 SATLKQOSTLAOMYPLQEIONTLVKLOALQONGSSVSEKSRMTIINTASTIYST 111
 DB 388 -AFLKER-----YLLAIQNNSHIETTDVYANLKR-NNTYSKIVA- 425
 QY 112 GKVCNPDNPQECILLEPGLNETMANSLDYNERLWAMESRSEVGKOLPLIEEYVLEKNE 171
 DB 426 -----PVSTEL-----YSELYF-----GNSING-YEDPAVISPI 453
 QY 172 MARAHYEDYDYMWDGVEVNGVDYDYSRGQLIEDVETPEELKPLYEHLAAVRAKLM 231
 DB 454 LGSFNAGATFRFR-TGNFNK-----KKQLOKEIVH-----HYNNMNMNDL 495
 QY 232 NAYPSIISIGCLPAHLIGDMGRWTVLSLVPEFGKPNIDVTDAVDAQARIFK 291
 DB 496 ELSQLEAPLNSRNVILNN--NRIYVCLMLNP--KSDIDINDIFIGATFKLYLS 550
 QY 292 E--AEKFFVSGLPNMTG-----FWNSMLTDRN-----VOKA 324
 DB 551 EKHDSRIYFVNSMNYEFGSELKFLREISFEKTKIOTPTTEGIDSLFCPRITIKNI 610
 QY 325 VCHPTAMULGKDRILIMCTVIMDELTAHHEMGIOTYMAVA----- 368
 DB 611 ILKPTATKINSEMESE--TENMLNRPATI-REKWHIPKVIYLAFGNRLMLLMDKHI 666
 QY 369 -----AOPFLRNGANEGHEAVGEIMS-----LSATPEK-----HUK 401
 DB 667 IILKELKKHGRILLESFINESNNEMLEIVPLKYSILKQDSFLIPRNRKHPNNIK 726

QY 402 ----SIGLSPDQEDNETELNPLKQALITVGLPTLYLEKRWVKEGELPKDQWKK 458
 DB 727 DWSIHSLIPKTYQD-----NFIQDYLLFPIETELKYNNEINFFYIKFKED--EDFK- 777
 QY 459 WMEKREIVGVBPVHDEFTYCDPASLFFVNSDYSPIRY-----TITLQFOFOEALC 512
 DB 778 -LRLRE-----DEDY-----SQIYFIMKMDQCLINSELDYISIVYVP 817
 QY 513 QAAKHGRLKDISSTAGOKLPLNMLRKGSEPTLALENVGAKMNYRPLNTFEP 572
 DB 818 EYRIGGPHVIDIENFEMYSIL--SINIQSE-FKIPKEPIYA--ISIDFLDLE- 870
 QY 573 LEFTWLKQKNSFVGSWTDWSPYADQSIKVRISLSKALGDAYEMDNEMTLRGSVAYA 632
 DB 871 -----INKSKEELLNNA--EDLYRSD----- 892
 QY 633 MROYFLAKNQMILFGEDVRANIKPRISNFFVTAPKKNVSDIIP----- 678
 DB 893 IREY-----KNILAKL-----TJPKNDYELIKKEFPNLHEFLFKI 928
 QY 679 -----RTEVERKAIRMSRINDAF--RLDNSLEFLGTOP 711
 DB 929 SIENLKKTKLOKSLYTSRISRIIGSFIMKCN--RIFGINP 966

RESULT 14

US-08-630-916A-46
 Sequence 46, Application US/08630916A
 Patent No. 6011137

GENERAL INFORMATION:
 APPLICANT: Pirozzi, Gregorio
 APPLICANT: Kay, Brian K.
 TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
 NUMBER OF SEQUENCES: 124
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/630,916A
 FILING DATE: 03-APR-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MISROCK, S. LESLIE
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 1101-203
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 896-8864/9741
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 683 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-630-916A-46

Query Match 3.0%; Score 116; DB 3; Length 683;
 Best Local Similarity 18.3%; Pred. No. 0.031;
 Matches 141; Conservative 107; Mismatches 265; Indels 258; Gaps 36;


```

QY 447 KGRIPKDOMKKMKWEKREIVGVEPVPHDEYCDPASLPHVSDYSFIRY-----TR 500
Db 769 KED---EDFK--LRLIRE-----DEY-----SQTYSFKMKDYCLNS 804
QY 501 TLVQFOFOALCOAKHGEPLHKCDISNTEAGOKLFNMLRLKSEPMFLALENVGAKN 560
Db 805 ELIYISIVDYVPEYRKGPHVEDIENF--MTDSLDSINIIQSE--FKIPKEFYA--- 859
QY 561 MNVRPLNTEPEPLFTWLKDONKNSFVGSWTDSPADOSIKYRISLSKALGKAYEMNDN 620
Db 860 ISIDFLDYLE-----INKEKEELLINNA--EDLYRSDN-- 892
QY 621 EMILFSSVAYAMQYFLKYNOMILFGEEDVRVANLKPRISNFTVAPKAVSDIIP-- 678
Db 893 -----IREY-----KNLAKL-----TPKNDYEILKKE 916
QY 679 -----RTEVEKAIMRSRINDAF-----RLNDSLEFLGIOP 711
Db 917 FPNLHEFLFKISILENKLKTKLOKSLYTSRSRIIGSFIMRCN-----RIFGINP 966

RESULT 16
US-08-466-961A-20
; Sequence 20, Application US/08466961A
; Patent No. 5843709
; GENERAL INFORMATION:
; APPLICANT: Entlian, Karl-Dieter
; APPLICANT: Gtz, Friedrich
; APPLICANT: Schnell, No. 5843709bert
; APPLICANT: Augustin, Johannes
; APPLICANT: Engelke, Gerhard
; APPLICANT: Rosenstein, Ralf
; APPLICANT: Kaletta, Corinna
; APPLICANT: Klein, Cora
; APPLICANT: Wieland, Bernd
; APPLICANT: Kupke, Thomas
; APPLICANT: Jung, G nther
; APPLICANT: Kellner, Roland
; TITLE OF INVENTION: Biosynthetic Process for the Preparation of
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/392,625
; FILING DATE: 22-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/876,791
; FILING DATE: 30-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/784,234
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0980004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

```

```

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 990 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-466-961A-20

Query Match          3.0% Score 115; DB 2; Length 990;
Best Local Similarity 17.9%; Pred. No. 0.071;
Matches 149; Conservative 116; Mismatches 235; Indels 334; Gaps 41;

2 TIEQAKTFLDKFNHE-----AEDL--FYOSLSAMWNTNITEENQNNMAGDKW 51
Db 343 TIRNYHEPFMDKYGPEQLVNLKOLLSDINGRGYKPKOSYSSNNI----- 387
QY 52 SAFLEKQSTLAOMPILOIOMLYKLOLQLOQNGSSVLSKSKRLNTILNTSTYST 111
Db 388 -AFLEK-----YLAIAONSHLEITENDVKNLK--NNTVSKINA-- 425
QY 112 GKVCNPDNPQCLLLEPLGELNEMANSIDYNERLWAMESRSEVGKOLRPLYEEVYLKNE 171
Db 426 -----PYSTEI-----YSEIY-----GNSIKG--YEDFAVISPI 453
QY 172 MARANHYEDYGDYKRGDYEVNGVDGYDSRGQLIEDVEHFEELKPLYEHLHAYRAKIM 231
Db 454 LGSEFAGATFGRP--TGNENIK-----KKNQLOKEIVH-----HYNNMNMENDL 495
QY 232 NAYPSYISPICCLRAHLIGDMGKRFNWLNSLYVPEGOKPRIDYTDAMDQAMDQRIEK 291
Db 496 EISOLNEAPLNSRVNLIINN--NRIYNTCLNLNP--KSDIDINDIFIGATFNKLYLS 550
QY 292 E--AEKEFVSGLPNNMTQG-----FWENSMILTDGNN-----VOKA 324
Db 551 EKHDSRIYFVGSNMFNENFEGSELYKFLREISFEKTKTQPTTEGIDSLPPCPRIITYKI 610
QY 325 VCHPTAMDGK-----GDFRIIMCTRYTMDDEFLTAH 355
Db 611 ILKPATWKINSEMESETEMNLRPATIRKWHIPADVIILARGDNRLIL-----NLNDK 663
QY 356 H-----EMGHIODMAVAAOPFLIRANGANGCHEANGELMS-----LSAAT 396
Db 664 HLIILKELEKKHGRIR-----TLESFINESNNERMLETVTPLYKRTSLKEOSFI 713
QY 397 PK-----HLK-----SIGLSPDEQNETEINFLKQALTIYGTLPFTYMLEKRWAVF 446
Db 714 PKNRKHPNNLKDWFSLHSIPKTYQD-----NFIDYLLPPTTELKVNNEFINKPFYIKF 768
QY 447 KGRIPKDOMKKMKWEKREIVGVEPVPHDEYCDPASLPHVSDYSFIRY-----TR 500
Db 769 KED---EDFK--LRLIRE-----DEY-----SQTYSFKMKDYCLNS 804
QY 501 TLVQFOFOALCOAKHGEPLHKCDISNTEAGOKLFNMLRLKSEPMFLALENVGAKN 560
Db 805 ELIYISIVDYVPEYRKGPHVEDIENF--MTDSLDSINIIQSE--FKIPKEFYA--- 859
QY 561 MNVRPLNTEPEPLFTWLKDONKNSFVGSWTDSPADOSIKYRISLSKALGKAYEMNDN 620
Db 860 ISIDFLDYLE-----INKEKEELLINNA--EDLYRSDN-- 892
QY 621 EMILFSSVAYAMQYFLKYNOMILFGEEDVRVANLKPRISNFTVAPKAVSDIIP-- 678
Db 893 -----IREY-----KNLAKL-----TPKNDYEILKKE 916
QY 679 -----RTEVEKAIMRSRINDAF-----RLNDSLEFLGIOP 711
Db 917 FPNLHEFLFKISILENKLKTKLOKSLYTSRSRIIGSFIMRCN-----RIFGINP 966

RESULT 17
US-08-844-059-2
; Sequence 2, Application US/08844059
; Patent No. 6001601
; GENERAL INFORMATION:

```

APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6001601el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19046
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,059
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607999.1
FILING DATE: 16-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-844-059-2

Query Match 2.9%; Score 113.5; DB 3; Length 665;
Best Local Similarity 19.2%; Pred. No. 0.052;
Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

QY 125 LLEPGINELMANSIDYNERIAMESRSEYKQRLPYEEVYVKNEMARANHEDYGY 184
DB 85 LLDISYDKFRTDDYHEKVA-----QVFERLLAQDDIY--LGEY 123
QY 185 WRGDEYVNGVDGYDYSRGOLIE-----DVHTFEERIKPLYEHLNAVY 226
DB 124 -SGMTYSV--DEEFFESQLAIEFRDAGNVTGGLAPSGHEHVMWSESEYFLRLSKYQR 180
QY 227 RAKIMNAVPSYISPIGCLPAHL-----LGDMMGRFTNLSLVPGOKPNIDY-TDA 278
DB 181 LVFEFFRAHPETIPDGLNMLNFIPEGLDLAVSRFTMGVPPVSPMKHVVYWIDA 240
QY 279 MV-----DQAWDA--QRTFKAEK-----FVSGLPN--MQ 307
DB 241 LLNYATALGYADQEHGDFKFMGTVEHMGKDLRFHSHYPLMLMDVYKLPRLIAH 300
QY 308 GFV---ENSMULTOPGNVQKAVCPFTAMDLAGKDFRILMCTKVMDPFLAHNEMGIQYD 364
DB 301 GNVVMDCKMSKSKGNVYVPEMLVERYGLDPLRYILMRNLVPVSDGTFPEDEYGLINTE 360
QY 365 MAYAAPFLR-----NGANEGFHEAVGEIMSLSATPKHLKSIGLSPFQEDNE 415
DB 361 LANDIGMLNRYVSMINKYFDGQIPAYVEGVTEFDHVALEVAE--KSID---DFTHME 414
QY 416 TEINFLKQALLTYGIL---PPTYMLEKRWMYFKGEIPRQDQ---MKKWKMKRRIYGV 469
DB 415 A-VDY--PRALIEVWTLISRTNKYIDETAPWVLDKDEALDQLASVSHQASIRYVAHL 471
QY 470 VEPVPHDEYCDPASLFEHVSNDYSEIRYVYRTLYQFOFQALCOAKNHGSLPKCDISNS 529
DB 472 IEP-----FMETSRAV----- 483

QY 530 TEAGOKLFNMLRCKSPFWTLAENV-----VGAKNMVYPLNFEPLFTWLD 579
DB 484 -----LQJILEE--VSSLNLSLADFPADVVAVGTPIFPLNNEEEI-AVIKE 531
QY 580 QNKSFWGWSIDWSP-----YAD-QSIRKVIS-----IKSALGDYAEW--- 617
DB 532 QMEGNKPAVEKEWNPDEVELKLNKDEIKFEDEKVEIRVAEYKESVKEGSDKLOFRD 591
QY 618 -NDNEMYLFRSSVAVARQYFLKYNQMLFGEEDVAVANLKPRISEFVTA PKNVSDI 676
DB 592 AGDEKROILSGIA-----KTY--PNDELVGKVVQVYVANKR----- 628
QY 677 IPRTVEKAIMSRKINDAERL--NDNSLEFLGIPTLGPV 717
DB 629 -----KMKKVVYSGMILSAEHDGKRTLLTVDPV--PN 660

RESULT 18
US-09-431-202-2
Sequence 2, Application US/09431202
Patent No. 6294175
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6294175el Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19046
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,202
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,059
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31456
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-431-202-2

Query Match 2.9%; Score 113.5; DB 4; Length 665;
Best Local Similarity 19.2%; Pred. No. 0.052;
Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

QY 125 LLEPGINELMANSIDYNERIAMESRSEYKQRLPYEEVYVKNEMARANHEDYGY 184
DB 85 LLDISYDKFRTDDYHEKVA-----QVFERLLAQDDIY--LGEY 123
QY 185 WRGDEYVNGVDGYDYSRGOLIE-----DVHTFEERIKPLYEHLNAVY 226

```

Db 124 -SGWTSVS--DEEFFTESQLAEVFRDEAGNTGGAIPSGHEVWVSESEYELRLSKYODR 180
QY 227 RAKIMNAPSYISPIGCLPAHL-----LGDMMRGREYNTLSLTVPGOKPNIDY-TDA 278
Db 181 LVEFFKAHEFTTPDGRLENMLRNIEGLEDIAKSRFTTFWGPVPSNPNHVVYVWIDA 240
QY 279 MV-----DQAWDA---ORIRKEAKR-----FVSGLPN--MTQ 307
Db 241 LLNATATAGYADEHGNDEKFMNGTVFHMWKGDLIRFHSIYWPILLMMDLVKLPDRLLAH 300
QY 308 GFW--ENSMLTDGPNQKAVCHPTAMDIGKDRILMCTVTMDLTLAHHEGHQYD 364
Db 301 GMFWKQCKMKSCKGNVYPPMLVERGLDPLRIYLMKNLPVSGDGTFFPEDYVGRINYE 360
QY 365 MAYAAQPELR-----NGANEGRHVAEGLMSLSAATPKHLKSLGILSPDPODNE 415
Db 361 LANLGNLLNRTVSMINKREYDQIPAYVEGTEFDHVAEAE--KSLA---DFHTME 414
QY 416 TEINELLKQALTYIGTL--PFTYMLKRWKRWYKGLIPKDW--MKKWMKREIIVG 469
Db 415 A-VUY--PRLEAVWTLISRTNKYIDETAPWVLDKDLALDQLASVSHQASIRVVAHL 471
QY 470 VEPVPHDEYCDPASLFHVSNDSYFIRYTRTYQFOFQALCOAAKHGEPHLKCDISNS 529
Db 472 IEP-----FMMETSRAY-----483
QY 530 TEAGQKLFNMLRLCKSEPTLALENV-----VGAKNMVRLPLNTEFLEPLTWLD 579
Db 484 -----LTQIGLEB--VSSLENTSLADPADVYVAKGTPLEPLRNLNEEL--AYIKE 531
QY 580 QNKSFWGSTMDSPP-----YAD-QSIVKRIS-----LKSALGDAYEY---617
Db 532 QMEGNKRAYEKENPDEVETKLKDELKEFEDPKVEIRVAEVEVSVEGSDLLQFRID 591
QY 618 -NDNEMTLFSSVAYANROYFLKYKNOMILFGEDEVANILKPRISNFEVTAAPKNVSDI 676
Db 592 AGGEDRQLISGIA---KYY---PNQDELVGKKQVYANLKNR-----628
QY 677 IPRTEVEKATIRMSRIRINDAFRL--NDNSLEFLGIQPIFGPPN 717
Db 629 -----KMMKKYVSGMILSAEHDKLTLLTVDPAN--PN 660

RESULT 19
US-08-491-357-2
; Sequence 2, Application US/08491357
; Patent No. 5716782
GENERAL INFORMATION:
APPLICANT: Golemis, Erica A.
APPLICANT: Law, Susan
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
TITLE OF INVENTION: ALTERATIONS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,357
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.

```

```

REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHEITICAL: NO
ANTI-SENSE: NO
US-08-491-357-2

Query Match
Best Local Similarity 2.7%; Score 104; DB 1; Length 834;
Matches 88; Conservative 59; Mismatches 142; Indels 140; Gaps 19;

QY 38 EENVQNMNNAQDKWSAFKE-----QSTLAQMTPIQE-IQNTLVKIQALQALQNGSSVYS 91
Db 442 ERHINEIRTVADKVELLEFLKELYHFKGAVANAAQPELLIHNMKREIQRV-EDSHQILS 500
QY 92 E-----DKSKRLNTLNT-----104
Db 501 QTSHDINECSWSLNLIAINKPQKCDLDRFYVAKTYVDKAKQLTTITINIAELFRPG 560
QY 105 -----NSTIY---STGKVCNP-DNPOEC--ILLEBGLNEIYA--NSIDY 140
Db 561 PGLHLKNGPESIMNSTEYRPHGSGOQLLHPGDHKAQAHNKLPLGLSKGAPDCSSDG 620
QY 141 NERLWAME-----SWRSVQKQLRPIVEEYVVLNENARAHYDYQDYRGDIYEVGV 194
Db 621 SERSMDDVDYHHLQKEFEFEQOKELLEKENIMKQNKQLEHHQ-----665
QY 195 DQYDYSKGLIDVHTFEERKPLYEHLHAYVRAKIMNAPSYISPIGCLPAHLIGDWG 254
Db 666 -----LSQFOLLBQ-----EITKPVENDISKW--KPSQSLPTNSGVAQDRQLLCFYTD 713
QY 255 REFTNLSYLVTFPGQKPNIDVDAMVDAMDAQ--RIKREKEF-----FVSGLPNM 305
Db 714 QCEITHRISL-----LNAIDALFSCVSSAQPRIYFVANSKFVILSAHKLVEIG-DTL 763
QY 306 TQGFWENSNLDPGVOKAVCHPTAMDIGKDRILMCTKYTMDQF--LTAHHEGHQY 363
Db 764 TRQVTAQDIRKNKVMNNSQGLCEQL-----TIVATKAAHLHPSTLQEKVHOVT 815
QY 364 DMAYAAQEP 372
Db 816 DISRNAQLE 824

RESULT 20
US-08-968-633-2
; Sequence 2, Application US/08968633
; Patent No. 6100384
GENERAL INFORMATION:
APPLICANT: Golemis, Erica A.
APPLICANT: Law, Susan
TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
TITLE OF INVENTION: ALTERATIONS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,633
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/491,357
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6100384 Relevant
TOPOLOGY: No. 6100384 Relevant
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-968-633-2

Query Match 2.7%; Score 104; DB 3; Length 834;

Best Local Similarity 20.5%; Pred. No. 0.61; Mismatches 142; Indels 140; Gaps 19;

Db 442 ERHINERTRAVDELVELEFHYFGAVANNAACLPETLHNKRELQRY-EDSHQILS 500
QY 92 E-----D-----DKRLNTILNT----- 104
Db 501 QTSHDINECSWSLNTILAIKPKKCDLDRFYVAVATVPDPAKOLTTTINTNAEALFRPG 560
QY 105 -----MSTY-----STKVCNP-DNPOEC--LLEPGLINEIMA--NSLDY 140
Db 561 PGSLHLKNGPESTINSTEYFHGSGOGLHPGDHKAQAHNKALPGLSKQAPDCSSDQ 620
QY 141 NERLAME-----SWREVGKQLRPLYEYVVLKNEANAHYEDYGDYWRGDEYNGV 194
Db 621 SERGMDYDYVHLQGEERERQKELKEKINIKOKMKOLEHHQ----- 665
QY 195 DGYYSRGOLIEDEHFEELIKPLYEHLHAYVRAKLMNAVPSYISPIGCLPAHLGDMWG 254
Db 666 -----LSOFQLEQ-----ETTKPEVDISKW---KPSQSLPTNSGVSAQDRQLCEYID 713
QY 255 RFTNTLSLVFPGQKPNIDVTDAVDAQ--RIKFAERF-----FVSVGLPNM 305
Db 714 QCEPHFTSL-----LNAIDLFCVSSAOPRIFVAHSKFVILSAHKLVEIG-DTL 763
QY 306 TQGWENSMULTPCGNVAKAACHPTAMDLGKDFILMCTKVYMDP--LTAHHEMGHIQY 363
Db 764 TROYTAODIRNKVWNSNOLCEOLK-----TIVATKMAALHPSTTALQEMVHOYT 815
QY 364 DMAYAAQPF 372
Db 816 DISRNAQLF 824

Search completed: May 26, 2003, 17:58:18
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:54:24 ; Search time 39 Seconds

(without alignments)
1830.880 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 1 STIEEQAKFTLDFKFNHEAD.....DNSLFLGTPPLGPPNDPP 720

Scoring table: BLOSUM62

Gapopen 10.0, Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3869	100.0	805	US-09-978-385-2	Sequence 2, Appl1
2	3869	100.0	805	US-10-158-847-142	Sequence 142, App
3	3865	99.9	805	US-10-114-893-86	Sequence 86, Appl
4	3620	93.6	681	US-09-968-384-25	Sequence 25, Appl
5	3620	93.6	681	US-10-138-847-140	Sequence 140, Appl
6	3620	93.6	711	US-09-969-384-13	Sequence 13, Appl
7	3620	93.6	711	US-10-158-847-138	Sequence 138, App
8	3251	84.0	805	US-09-978-385-6	Sequence 6, Appl1
9	3233	83.6	805	US-09-978-385-9	Sequence 9, Appl1
10	2897	74.9	555	US-10-028-072-72	Sequence 72, Appl
11	2897	74.9	555	US-10-121-049-72	Sequence 72, Appl
12	2897	74.9	555	US-10-123-904-72	Sequence 72, Appl
13	2897	74.9	555	US-10-140-470-72	Sequence 72, Appl
14	2897	74.9	555	US-10-175-746-72	Sequence 72, Appl
15	2897	74.9	555	US-10-176-918-72	Sequence 72, Appl
16	2897	74.9	555	US-10-137-865-72	Sequence 72, Appl
17	2897	74.9	555	US-10-140-474-72	Sequence 72, Appl
18	2897	74.9	555	US-10-142-431-72	Sequence 72, Appl
19	2897	74.9	555	US-10-142-431-72	Sequence 72, Appl

20	2897	74.9	555	US-10-143-114-72	Sequence 72, Appl
21	2897	74.9	555	US-10-140-002-72	Sequence 72, Appl
22	2897	74.9	555	US-10-142-419-72	Sequence 72, Appl
23	2897	74.9	555	US-10-123-262-72	Sequence 72, Appl
24	2897	74.9	555	US-10-121-050-72	Sequence 72, Appl
25	2897	74.9	555	US-10-141-755-72	Sequence 72, Appl
26	2897	74.9	555	US-10-143-037-72	Sequence 72, Appl
27	2897	74.9	555	US-10-123-108-72	Sequence 72, Appl
28	2897	74.9	555	US-10-123-236-72	Sequence 72, Appl
29	2897	74.9	555	US-10-123-261-72	Sequence 72, Appl
30	2897	74.9	555	US-10-140-921-72	Sequence 72, Appl
31	2897	74.9	555	US-10-140-928-72	Sequence 72, Appl
32	2897	74.9	555	US-10-121-045-72	Sequence 72, Appl
33	2897	74.9	555	US-10-123-293-72	Sequence 72, Appl
34	2897	74.9	555	US-10-123-903-72	Sequence 72, Appl
35	2897	74.9	555	US-10-124-819-72	Sequence 72, Appl
36	2897	74.9	555	US-10-124-822-72	Sequence 72, Appl
37	2897	74.9	555	US-10-140-925-72	Sequence 72, Appl
38	2897	74.9	555	US-10-160-498-72	Sequence 72, Appl
39	2897	74.9	555	US-10-121-041-72	Sequence 72, Appl
40	2897	74.9	555	US-10-121-043-72	Sequence 72, Appl
41	2897	74.9	555	US-10-121-047-72	Sequence 72, Appl
42	2897	74.9	555	US-10-123-215-72	Sequence 72, Appl
43	2897	74.9	555	US-10-123-902-72	Sequence 72, Appl
44	2897	74.9	555	US-10-123-902-72	Sequence 72, Appl
45	2897	74.9	555	US-10-123-908-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-978-385-2
Sequence 2, Application US/09978385
Patent No. US2002017221A1
GENERAL INFORMATION:
APPLICANT: Piddington, Christopher S.
APPLICANT: Petrie, Charles
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: ZINC2: A HUMAN METALLOENZYME
FILE REFERENCE: 99-24C1
CURRENT APPLICATION NUMBER: US/09/978,385
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 60/133,952
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: 60/151,181
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: 09/563,516
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 805
TYPE: PRT
ORGANISM: Homo sapiens

Query Match	100.0%	Score 3869;	DB 9;	Length 805;
Best Local Similarity	100.0%	Pred. No. 1.9e-312;		
Matches 720;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
DB	1	STIEEQAKFTLDFKFNHEADLFYSSSLASNNYNTNTEENVOMNNAAGKMSAFLEKOST	60	
DB	19	STIEEQAKFTLDFKFNHEADLFYSSSLASNNYNTNTEENVOMNNAAGKMSAFLEKOST	78	
DB	61	LAQMPPLQETONTLVKIQDALQONGSSVLSDEKSKRLNTIITMGSTVSTGKVCNPDP	120	
DB	79	LAQMPPLQETONTLVKIQDALQONGSSVLSDEKSKRLNTIITMGSTVSTGKVCNPDP	138	
DB	121	QECILLPEGLNETMANSIDYNERLMAWESRSEVGKOLPIVEEYVVLAKENARAHYD	180	
DB	139	QECILLPEGLNETMANSIDYNERLMAWESRSEVGKOLPIVEEYVVLAKENARAHYD	198	

QY 181 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHTEFEIKPLYEHLAAVYRAKLMANAPSYISP 240
 DB 199 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHTEFEIKPLYEHLAAVYRAKLMANAPSYISP 258
 QY 241 IGCPLAHLGLDMGGRFTNLYSLTVPGQKPNIDYTDAMVDQAMDQORIFKEAEKFEVSV 300
 DB 259 IGCPLAHLGLDMGGRFTNLYSLTVPGQKPNIDYTDAMVDQAMDQORIFKEAEKFEVSV 318
 QY 301 GLPNMTQGFENSMUTDPGNVQKAVCHPTAMDGLGKDFRILMCTKYTMDDFLTAHHEMGH 360
 DB 319 GLPNMTQGFENSMUTDPGNVQKAVCHPTAMDGLGKDFRILMCTKYTMDDFLTAHHEMGH 378
 QY 361 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKISGLISPDFOEDNETEINF 420
 DB 379 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKISGLISPDFOEDNETEINF 438
 QY 421 LKQALITVGLPTFTYMLEKRWMMVFKEIIPKQDMKKWEMKREITGVVEPVPHDETYC 480
 DB 439 LKQALITVGLPTFTYMLEKRWMMVFKEIIPKQDMKKWEMKREITGVVEPVPHDETYC 498
 QY 481 DPASLFHVSNDYSFIRYTRITLYQFOFQALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
 DB 499 DPASLFHVSNDYSFIRYTRITLYQFOFQALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
 QY 541 RLGRSEPTTALLENVGAKNMNVRLNTEFEPLFTWLKDONKNSFVGWSTDMSPYADQSI 600
 DB 559 RLGRSEPTTALLENVGAKNMNVRLNTEFEPLFTWLKDONKNSFVGWSTDMSPYADQSI 618
 QY 601 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 660
 DB 619 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 678
 QY 661 ISFNEFVAPKAVSDIIPRTEVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNNQP 720
 DB 679 ISFNEFVAPKAVSDIIPRTEVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNNQP 738

RESULT 2

US-10-158-847-142
 ; Sequence 142, Application US/10158847
 ; Publication No. US2003009157A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tom Parry et al.
 ; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
 ; FILE REFERENCE: P557
 ; CURRENT APPLICATION NUMBER: US/10/158,847
 ; CURRENT FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: 60/295,004
 ; PRIOR FILING DATE: 2001-06-04
 ; NUMBER OF SEQ ID NOS: 158
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 142
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-158-847-142

Query Match 100.0%; Score 3869; DB 9; Length 805;
 Best Local Similarity 100.0%; Pred. No. 1,9e-312;
 Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQAKTELDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKWSAFLEQST 60
 DB 19 STIEQAKTELDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKWSAFLEQST 78
 QY 61 LAQWYPIQOEIONTLVKIQLOALOQNGSSVLESDSKRLNTLIMTSTVSTGVQCPNDP 120
 DB 79 LAQWYPIQOEIONTLVKIQLOALOQNGSSVLESDSKRLNTLIMTSTVSTGVQCPNDP 138
 QY 121 QECLELPGLEINANSIDYERLMAWESWSEVQKOLRPLDEEYVVLKKNEMARANYED 180
 DB 139 QECLELPGLEINANSIDYERLMAWESWSEVQKOLRPLDEEYVVLKKNEMARANYED 198

QY 181 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHTEFEIKPLYEHLAAVYRAKLMANAPSYISP 240
 DB 199 YGDYWRGDEYVNGVGYDYSRGQLIEDVEHTEFEIKPLYEHLAAVYRAKLMANAPSYISP 258
 QY 241 IGCPLAHLGLDMGGRFTNLYSLTVPGQKPNIDYTDAMVDQAMDQORIFKEAEKFEVSV 300
 DB 259 IGCPLAHLGLDMGGRFTNLYSLTVPGQKPNIDYTDAMVDQAMDQORIFKEAEKFEVSV 318
 QY 301 GLPNMTQGFENSMUTDPGNVQKAVCHPTAMDGLGKDFRILMCTKYTMDDFLTAHHEMGH 360
 DB 319 GLPNMTQGFENSMUTDPGNVQKAVCHPTAMDGLGKDFRILMCTKYTMDDFLTAHHEMGH 378
 QY 361 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKISGLISPDFOEDNETEINF 420
 DB 379 IQYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKISGLISPDFOEDNETEINF 438
 QY 421 LKQALITVGLPTFTYMLEKRWMMVFKEIIPKQDMKKWEMKREITGVVEPVPHDETYC 480
 DB 439 LKQALITVGLPTFTYMLEKRWMMVFKEIIPKQDMKKWEMKREITGVVEPVPHDETYC 498
 QY 481 DPASLFHVSNDYSFIRYTRITLYQFOFQALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
 DB 499 DPASLFHVSNDYSFIRYTRITLYQFOFQALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
 QY 541 RLGRSEPTTALLENVGAKNMNVRLNTEFEPLFTWLKDONKNSFVGWSTDMSPYADQSI 600
 DB 559 RLGRSEPTTALLENVGAKNMNVRLNTEFEPLFTWLKDONKNSFVGWSTDMSPYADQSI 618
 QY 601 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 660
 DB 619 KVRISLSALGDKAYEMNDNEMTLFRSSVAYAMQYFLKYNOMILFGEEDVAVANLKR 678
 QY 661 ISFNEFVAPKAVSDIIPRTEVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNNQP 720
 DB 679 ISFNEFVAPKAVSDIIPRTEVEKAIKMSRSRINDAFLRNDNSLEFLGIQPTLGPNNQP 738

RESULT 3

US-10-114-893-86
 ; Sequence 86, Application US/10114893
 ; Publication No. US20020193567A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallee, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Bowman, Michael R.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Carlin-Duckett, McKeough
 ; APPLICANT: Kelleher, Kerry S.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: GI 6000-10A
 ; CURRENT APPLICATION NUMBER: US/10/114,893
 ; EARLIER APPLICATION NUMBER: 09/413,232
 ; EARLIER FILING DATE: 1999-10-06
 ; NUMBER OF SEQ ID NOS: 321
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 86
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-114-893-86

Query Match 99.9%; Score 3865; DB 9; Length 805;
 Best Local Similarity 99.9%; Pred. No. 4.1e-312;
 Matches 719; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 STIEQAKTFIDKRNHEADLFYQSSLASMYNTNITEENVQNNANNGDKMSAFLEQST 60
DB 19 STIEQAKTFIDKRNHEADLFYQSSLASMYNTNITEENVQNNANNGDKMSAFLEQST 78
OY 61 LAQWYPLQEIQNLVVKLOLQALQONGSSVLSSEKSKRLNTLNTMTSTTSTGKVCNPNP 120
DB 79 LAQWYPLQEIQNLVVKLOLQALQONGSSVLSSEKSKRLNTLNTMTSTTSTGKVCNPNP 138
OY 121 QECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYEYVYLKKNEMARANYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYEYVYLKKNEMARANYED 198
OY 181 YGDWVRGDEYVNGVDGYSRGQLEIEVEHTFEETIKLYEHLAAYRAKIMNAPSYISP 240
DB 199 YGDWVRGDEYVNGVDGYSRGQLEIEVEHTFEETIKLYEHLAAYRAKIMNAPSYISP 258
OY 241 IGCPLPAHLGDMGREFNTNLSLVPPGQKPNIDVTAMDQAMDQORIKKEKEKFFVSY 300
DB 259 IGCPLPAHLGDMGREFNTNLSLVPPGQKPNIDVTAMDQAMDQORIKKEKEKFFVSY 318
OY 301 GLPMTQGFENSKMLTDPGNVQKAVCHPTAMDLGKDPRLIMCTKYVMDPFLTAHHEMGH 360
DB 319 GLPMTQGFENSKMLTDPGNVQKAVCHPTAMDLGKDPRLIMCTKYVMDPFLTAHHEMGH 378
OY 361 IYQDMAVAAOPFLLRNGANGEGHEAVGETMSLSAATPKHLKSTIGLSPDEQEDNETEINF 420
DB 379 IYQDMAVAAOPFLLRNGANGEGHEAVGETMSLSAATPKHLKSTIGLSPDEQEDNETEINF 438
OY 421 LKQALTYVGTLPFTYMLEKRMWVFGELIPKQDMKKWMEKRELYGVVEPVPHEDETC 480
DB 439 LKQALTYVGTLPFTYMLEKRMWVFGELIPKQDMKKWMEKRELYGVVEPVPHEDETC 498
OY 481 DPASLHVSNDSYFIRYTRTYLYQFOFOALCOAAHCEGPLKCDISNSTEAGOKLFNML 540
DB 499 DPASLHVSNDSYFIRYTRTYLYQFOFOALCOAAHCEGPLKCDISNSTEAGOKLFNML 558
OY 541 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTLKDQNNKSPFGNSTDMSPYADOSI 600
DB 559 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTLKDQNNKSPFGNSTDMSPYADOSI 618
OY 601 KVAISLKSALGDKAYEMNDNEMTLFRSSVAYAMROYFLKVKNNOMILFGEDEVYANLKPR 660
DB 619 KVAISLKSALGDKAYEMNDNEMTLFRSSVAYAMROYFLKVKNNOMILFGEDEVYANLKPR 678
OY 661 ISFNEFTVAPKNVSDIIPRTVEKKAIRMSRSRINDAFRLNDSLEFLGIQPLGPNOQP 720
DB 679 ISFNEFTVAPKNVSDIIPRTVEKKAIRMSRSRINDAFRLNDSLEFLGIQPLGPNOQP 738

RESULT 4
US-09-969-384-25
: Sequence 25, Application US/09969384
: Publication No. US20020192749A1
: GENERAL INFORMATION:
: APPLICANT: Moore, et al.
: TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
: FILE REFERENCE: PT055P1
: CURRENT APPLICATION NUMBER: US/09/969,384
: PRIOR FILING DATE: 2001-10-03
: PRIOR APPLICATION NUMBER: PCT/US01/10542
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 60/236,384
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/194,118
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 25
: LENGTH: 681
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE

```

```

: LOCATION: (219)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (240)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (499)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-969-384-25

Query Match          93.6%; Score 3620; DB 9; Length 681;
Best Local Similarity 99.3%; Pred. No. 7e-292;
Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 44 MNNAGDKMSAFLEKEOSTLAQMTPLQEIQLVVKLOLQALQONGSSVLSSEKSKRLNTLNT 103
DB 1 MNNAGDKMSAFLEKEOSTLAQMTPLQEIQLVVKLOLQALQONGSSVLSSEKSKRLNTLNT 60
OY 104 TMSITSTGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 163
DB 61 TMSITSTGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWESRSEVQKQRLPYE 120
OY 164 EYVYLKKNEMARANYEDYGDWVRGDEYVNGVDGYSRGQLEIEVEHTFEETIKLYEHL 223
DB 121 EYVYLKKNEMARANYEDYGDWVRGDEYVNGVDGYSRGQLEIEVEHTFEETIKLYEHL 180
OY 224 AYVRKIMNAPSYISTPGCLPFAHLGDMGREFNTNLSLVPPGQKPNIDVTAMDQOA 283
DB 181 AYVRKIMNAPSYISTPGCLPFAHLGDMGREFNTNLSLVPPGQKPNIDVTAMDQOA 240
OY 284 WDAQRIFKEAEKEFFSVGLPNTQGFENSKMLTDPGNVQKAVCHPTAMDLGKDPRLIMC 343
DB 241 WDAQRIFKEAEKEFFSVGLPNTQGFENSKMLTDPGNVQKAVCHPTAMDLGKDPRLIMC 300
OY 344 TKVYMDPFLTAHHEMGHIQYDMAVAAOPFLLRNGANGEGHEAVGEINLSAATPKHLKSI 403
DB 301 TKVYMDPFLTAHHEMGHIQYDMAVAAOPFLLRNGANGEGHEAVGEINLSAATPKHLKSI 360
OY 404 GLSPDEQEDNETEINFLLKQALTYVGTLPFTYMLEKRMWVFGELIPKQDMKKWMEBK 463
DB 361 GLSPDEQEDNETEINFLLKQALTYVGTLPFTYMLEKRMWVFGELIPKQDMKKWMEBK 420
OY 464 REYGVVEPVPHEDETCDPASLHVSNDSYFIRYTRTYLYQFOFOALCOAAHCEGPLK 523
DB 421 REYGVVEPVPHEDETCDPASLHVSNDSYFIRYTRTYLYQFOFOALCOAAHCEGPLK 480
OY 524 CDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNMNVRPLNTEPELFTLKDQNNK 583
DB 481 CDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNMNVRPLNTEPELFTLKDQNNK 540
OY 584 SFVGMSTDMSPYADOSIKVAISLKSALGDKAYEMNDNEMTLFRSSVAYAMROYFLKVK 643
DB 541 SFVGMSTDMSPYADOSIKVAISLKSALGDKAYEMNDNEMTLFRSSVAYAMROYFLKVK 600
OY 644 MILFGEDEVYANLKPRISFNEFTVAPKNVSDIIPRTVEKKAIRMSRSRINDAFRLNDS 703
DB 601 MILFGEDEVYANLKPRISFNEFTVAPKNVSDIIPRTVEKKAIRMSRSRINDAFRLNDS 660
OY 704 LEFLGIQPLGPNOQP 720
DB 661 LEFLGIQPLGPNOQP 677

RESULT 5
US-10-158-847-140
: Sequence 140, Application US/10158847
: Publication No. US20030091557A1
: GENERAL INFORMATION:
: APPLICANT: Tom Parry et al.
: TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
: FILE REFERENCE: PF557
: CURRENT APPLICATION NUMBER: US/10/158,847
: CURRENT FILING DATE: 2002-06-03

```

PRIOR APPLICATION NUMBER: 60/295,004
 PRIOR FILING DATE: 2001-06-04
 NUMBER OF SEQ ID NOS: 158
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 140
 LENGTH: 681
 TYPE: PRT
 ORGANISM: homo sapiens
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (219)..(219)
 OTHER INFORMATION: xaa equals any amino acid
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (240)..(240)
 OTHER INFORMATION: xaa equals any amino acid
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (499)..(499)
 OTHER INFORMATION: xaa equals any amino acid
 US-10-158-847-140

Query Match 93.6%; Score 3620; DB 9; Length 681;
 Best Local Similarity 99.3%; Pred. No. 7e-292;

Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

44 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQNGSSVLSSEKSKRLNTLN 103
 1 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQNGSSVLSSEKSKRLNTLN 60
 104 TMSITSTGKVCNPNPOECILLPEGLNEIMANSIDYNERLWAMESRSEVGKQLRPLYE 163
 61 TMSITSTGKVCNPNPOECILLPEGLNEIMANSIDYNERLWAMESRSEVGKQLRPLYE 120
 164 EYVLAKNEMARANHYEDYDYGWDYEVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 223
 121 EYVLAKNEMARANHYEDYDYGWDYEVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 180
 224 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFNTNLXSLTVPGQKPNIDVTDAVDQA 283
 181 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFNTNLXSLTVPGQKPNIDVTDAVDQA 240
 284 WDAORIFKEAEKEFEYSVGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRLIMC 343
 241 WDAORIFKEAEKEFEYSVGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRLIMC 300
 344 TKTMDDELTAHHEMHIQYDMAVYAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 403
 301 TKTMDDELTAHHEMHIQYDMAVYAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 360
 404 GLSPDFQEDNTEINFLKQALITVGLTPFTYMLEKRMWVFKEIPKQDMKKWEMK 463
 361 GLSPDFQEDNTEINFLKQALITVGLTPFTYMLEKRMWVFKEIPKQDMKKWEMK 420
 464 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGEPPLK 523
 421 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGEPPLK 480
 524 CDISNSTEAGQKLFNNMLRLGKSEPTLLENVVGAKNMNVRPLNTEPELFTWLKDQKN 583
 481 CDISNSTEAGQKLFNNMLRLGKSEPTLLENVVGAKNMNVRPLNTEPELFTWLKDQKN 540
 584 SPVGMSTWSPYADQSIKVRISLSKALGDKATYEMNDENKTLFFSSVAYAMROFELVYKQ 643
 541 SPVGMSTWSPYADQSIKVRISLSKALGDKATYEMNDENKTLFFSSVAYAMROFELVYKQ 600
 644 MILEGEEDVYANAKPRISNFETAPKANSDIIPREVEKAIKMSRSRINDAFRLNDS 703
 601 MILEGEEDVYANAKPRISNFETAPKANSDIIPREVEKAIKMSRSRINDAFRLNDS 660
 704 LEFLGIQPTLGPNNQP 720
 661 LEFLGIQPTLGPNNQP 677

RESULT 6
 US-09-969-384-13
 Sequence 13, Application US/09969384
 Publication No. US20020192749A1
 GENERAL INFORMATION:
 APPLICANT: Moore, et al.
 TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
 FILE REFERENCE: P055P1
 CURRENT APPLICATION NUMBER: US/09/969,384
 PRIOR FILING DATE: 2001-10-03
 PRIOR APPLICATION NUMBER: PCT/US01/10542
 PRIOR FILING DATE: 2001-04-02
 PRIOR APPLICATION NUMBER: 60/236,384
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/194,118
 PRIOR FILING DATE: 2000-04-03
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 13
 LENGTH: 711
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (219)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (240)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (499)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 US-09-969-384-13

Query Match 93.6%; Score 3620; DB 9; Length 711;
 Best Local Similarity 99.3%; Pred. No. 7.5e-292;

Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

44 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQNGSSVLSSEKSKRLNTLN 103
 1 MNNAGDKMSAFLEKQSTLAQMYPLQEIONTLVKIQLOALQNGSSVLSSEKSKRLNTLN 60
 104 TMSITSTGKVCNPNPOECILLPEGLNEIMANSIDYNERLWAMESRSEVGKQLRPLYE 163
 61 TMSITSTGKVCNPNPOECILLPEGLNEIMANSIDYNERLWAMESRSEVGKQLRPLYE 120
 164 EYVLAKNEMARANHYEDYDYGWDYEVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 223
 121 EYVLAKNEMARANHYEDYDYGWDYEVNGVDYDYSRGQLIEDVEHTFEIKPLYEHLH 180
 224 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFNTNLXSLTVPGQKPNIDVTDAVDQA 283
 181 AYVRKLMANAYPSYISPIGCLPAHLGDMGRFNTNLXSLTVPGQKPNIDVTDAVDQA 240
 284 WDAORIFKEAEKEFEYSVGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRLIMC 343
 241 WDAORIFKEAEKEFEYSVGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRLIMC 300
 344 TKTMDDELTAHHEMHIQYDMAVYAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 403
 301 TKTMDDELTAHHEMHIQYDMAVYAAOPFLLRNGANEGFHEAVGEIMSLSAATPKHLKSI 360
 404 GLSPDFQEDNTEINFLKQALITVGLTPFTYMLEKRMWVFKEIPKQDMKKWEMK 463
 361 GLSPDFQEDNTEINFLKQALITVGLTPFTYMLEKRMWVFKEIPKQDMKKWEMK 420
 464 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGEPPLK 523
 421 RELVGVVEVPYPHDETCDPASLFHVSNDYSFIRYTRTLVYQFOFQALCOAAKHGEPPLK 480
 524 CDISNSTEAGQKLFNNMLRLGKSEPTLLENVVGAKNMNVRPLNTEPELFTWLKDQKN 583

Db 481 CDSINSTAGORLFLNMLRKGKSEPTWLTALENVGAKMNNVPLNTNFEPLTWLKDQKN 540
QY 584 SFVGNSTWSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 643
Db 541 SFVGNSTWSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 600
QY 644 MLLGGEEDVRANLKPRISENFVYAPKNSDIIIPTEVEKATIRMSRSRINDAFRLDMS 703
Db 601 MLLGGEEDVRANLKPRISENFVYAPKNSDIIIPTEVEKATIRMSRSRINDAFRLDMS 660
QY 704 LEFLGIPTLGGPPNPP 720
Db 661 LEFLGIPTLGGPPNPP 677

RESULT 7
US-10-158-847-138
; Sequence 138, Application US/10158847
; Publication No. US20030091557A1
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PE557
; CURRENT APPLICATION NUMBER: US/10/158,847
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 711
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (219)..(219)
; OTHER INFORMATION: xaa equals any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (240)..(240)
; OTHER INFORMATION: xaa equals any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (499)..(499)
; OTHER INFORMATION: xaa equals any amino acid
US-10-158-847-138

Query Match 93.6%; Score 3620; DB 9; Length 711;
Best Local Similarity 99.3%; Pred. No. 7, 5e-292;
Matches 672; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 44 MNNAGDKMSAFLEKQSTLAOMYPLQEIOMLTKYLOLQALQOQSSVLSSEKSKRLNTLN 103
Db 1 MNNAGDKMSAFLEKQSTLAOMYPLQEIOMLTKYLOLQALQOQSSVLSSEKSKRLNTLN 60
QY 104 TMSITISGKVCNPDPNOCILLLEPGLNEIMANSIDYNERIMAMESNSSEVGKOLRPLYE 163
Db 61 TMSITISGKVCNPDPNOCILLLEPGLNEIMANSIDYNERIMAMESNSSEVGKOLRPLYE 120
QY 164 EYVVLKNEKARAHNEDEGDYWRGDEYNGVDGDIYSGQLEIEVEHTEERIKPLYEHL 223
Db 121 EYVVLKNEKARAHNEDEGDYWRGDEYNGVDGDIYSGQLEIEVEHTEERIKPLYEHL 180
QY 224 AYVRAKLMAVPSYISPIGCLPAHLGDMGRFTNTLSLVPPGOKPNIDVDYDAMDOA 263
Db 181 AYVRAKLMAVPSYISPIGCLPAHLGDMGRFTNTLSLVPPGOKPNIDVDYDAMDOA 240
QY 284 WDAQIFFEAEKFFSVGLPMNTOGFENSLTDPGCVQAVACHPTAMDLGKGDFTILMC 343
Db 241 WDAQIFFEAEKFFSVGLPMNTOGFENSLTDPGCVQAVACHPTAMDLGKGDFTILMC 300
QY 344 TKVTMDFLTAHHEMGIQYDMAVAAPFLLRNGANGCFEAVGEIMSLSAATPKHLKSI 403

Db 301 TKVTMDFLTAHHEMGIQYDMAVAAPFLLRNGANGCFEAVGEIMSLSAATPKHLKSI 360
QY 404 GLSPDPODNTETENFLKQALTVGLPTTYMLEKRWVYFGEIIPKQOMKRWEMK 463
Db 361 GLSPDPODNTETENFLKQALTVGLPTTYMLEKRWVYFGEIIPKQOMKRWEMK 420
QY 464 REIVGVPEPPEHDETCBPASLFHSNDYSIRYRTYRTYPOFOQALCOAAHHEGLK 523
Db 421 REIVGVPEPPEHDETCBPASLFHSNDYSIRYRTYRTYPOFOQALCOAAHHEGLK 480
QY 524 CDSINSTAGORLFLNMLRKGKSEPTWLTALENVGAKMNNVPLNTNFEPLTWLKDQKN 583
Db 481 CDSINSTAGORLFLNMLRKGKSEPTWLTALENVGAKMNNVPLNTNFEPLTWLKDQKN 540
QY 584 SFVGNSTWSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 643
Db 541 SFVGNSTWSPYADOSIKVRSLSKALDRAKAYENDNEMTLFRSSVAYAMROYFLKVKNO 600
QY 644 MLLGGEEDVRANLKPRISENFVYAPKNSDIIIPTEVEKATIRMSRSRINDAFRLDMS 703
Db 601 MLLGGEEDVRANLKPRISENFVYAPKNSDIIIPTEVEKATIRMSRSRINDAFRLDMS 660
QY 704 LEFLGIPTLGGPPNPP 720
Db 661 LEFLGIPTLGGPPNPP 677

RESULT 8
US-09-978-385-6
; Sequence 6, Application US/09978385
; Patent No. US2002017211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE1: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Mouse
US-09-978-385-6

Query Match 84.0%; Score 3251; DB 9; Length 805;
Best Local Similarity 83.1%; Pred. No. 3, 8e-261;
Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

QY 1 STIEBQKTEFLKHEHAEFLFYOSLSASNNYNTNITEENQNNMNGDKMSAFLEKQST 60
Db 19 STIEBQKTEFLKHEHAEFLFYOSLSASNNYNTNITEENQNNMNGDKMSAFLEKQST 78
QY 61 LAOMYPLQEIOMLTKYLOLQALQOQSSVLSSEKSKRLNTLNMTSITISGVCNPN 120
Db 79 TAQSFSLQEIOTPIIRKQLQALQOQSSVLSADKNKQNTLNTMTSITISGVCNPN 138
QY 121 QCLLEPGLNEIMANSIDYNERIMAMESNSSEVGKOLRPLYEYVVLKNEKARAHNEDE 180
Db 139 QCLLEPGLNEIMANSIDYNERIMAMESNSSEVGKOLRPLYEYVVLKNEKARAHNEDE 198
QY 181 YGDYWRGDEYNGVDGDIYSRQGLEIEVEHTEERIKPLYEHLTAAYPAKMAVPSYIS 240

Db 199 YGDYWRGDYEAEGADGYNRNQLLIEDVETFAEIKPLYEHLHAYVRKLMJTYSTSP 258
 QY 241 IGCIPAHLLDGMGREFWNTNLSLTPVPGOKPNIDVDAVDAQMPAQRIFKEAEKFPVS 300
 Db 259 TGCIPAHLLDGMGREFWNTNLSLTPVPAOKPNIDVDAVDAQMPAQRIFKEAEKFPVS 318
 QY 301 GLPMTGOFWENSMITDPGNVQAVCHPTAMDLGKDFRLMCTKYTMDFLTAHHEGH 360
 Db 319 GLPMTGOFWENSMITDPGNVQAVCHPTAMDLGKDFRLMCTKYTMDFLTAHHEGH 378
 QY 361 IGYMAYRAOPEFLLRNANGFHEAVGEIMSLNATPKHLKSGLLPSDFODESEIINF 420
 Db 379 IGYMAYRAOPEFLLRNANGFHEAVGEIMSLNATPKHLKSGLLPSDFODESEIINF 438
 QY 421 LKQALITVGLPPTMYLEKRWAVFKEGEPKQDMKKWEMKREIVGVEPVPDETC 480
 Db 439 LKQALITVGLPPTMYLEKRWAVFKEGEPKQDMKKWEMKREIVGVEPVPDETC 498
 QY 481 DPASLEHVSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
 Db 499 DPASLEHVSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
 QY 541 RIGKSEPTLLENVYGAKNANVRPLNTEPEFTWMLDKONKSEVGMSTWSPYADQSI 600
 Db 559 SIGNSEPTLLENVYGAKNANVRPLNTEPEFTWMLDKONKSEVGMSTWSPYADQSI 618
 QY 601 KVRISLKSALGDKAYEMNDENKYLFRSSVAYAMQOYLKYNQMLLGEEDVAVANLKR 660
 Db 619 KVRISLKSALGDKAYEMNDENKYLFRSSVAYAMQOYLKYNQMLLGEEDVAVANLKR 678
 QY 661 ISFNFPVAPKVNDSIIPRTVEYKARMSRSRINDAFRLNDSLEFLGIOTPLGPPNPQ 720
 Db 679 VSFFYFVAPKVNDSIIPRTVEYKARMSRSRINDAFRLNDSLEFLGIOTPLGPPNPQ 738

RESULT 9

US-09-978-385-9
 ; Sequence 9, Application US/09978385
 ; Patent No. US2002017211A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Piddington, Christopher S.
 ; APPLICANT: Petrie, Charles
 ; APPLICANT: Shoemaker, Kimberly E.
 ; APPLICANT: Bishop, Paul D.
 ; TITLE OR INVENTION: ZACE2: A HUMAN METALLOENZYME
 ; FILE REFERENCE: 99-24C1
 ; CURRENT APPLICATION NUMBER: US/09/978,385
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 60/133,952
 ; PRIOR FILING DATE: 1999-05-13
 ; PRIOR APPLICATION NUMBER: 60/151,181
 ; PRIOR FILING DATE: 1999-08-27
 ; PRIOR APPLICATION NUMBER: 09/563,516
 ; PRIOR FILING DATE: 2000-05-03
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: fastseq for Windows Version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 805
 ; TYPE: PRT
 ; ORGANISM: Mouse
 ; US-09-978-385-9

Query Match 83.6%; Score 3233; DB 9; Length 805;
 Best Local Similarity 82.8%; Pred. No. 1,2e-259;
 Matches 596; Conservative 49; Mismatches 75; Indels 0; Gaps 0;

QY 1 STIEQAKTFLDKRNHAEELFYOSSLASMYNTNTTEENVOMNNAAGDKMSAFLEKQST 60
 Db 19 SLTEENAKTFLNNNOEAEDISTOSSLASMYNTNTTEENAOXMSAAKMSAFLEKQSK 78
 QY 61 LAOMYPLQEIQNTLVKQLQALQONGSSVLSSEKSKRLNTLITMSTIYSTGKVCNPNP 120
 Db 79 TAOFSQEIOTPIIKKQALQALQONGSSSALSADKNKQNTLITMSTIYSTGKVCNPNP 138

QY 121 QECILLPEGNEITANSIDYNERLNAWESRSEVGKOLRPLYEEYVLKNEKARAHYED 180
 Db 139 QECILLPEGNEITANSIDYNERLNAWESRSEVGKOLRPLYEEYVLKNEKARAHYED 198
 QY 181 YGDYWRGDYEVNGVDYDYSRGOLLIEDVETFEIKPLYEHLHAYVRKLMJTYSTSP 240
 Db 199 YGDYWRGDYEAEGADGYNRNQLLIEDVETFAEIKPLYEHLHAYVRKLMJTYSTSP 258
 QY 241 IGCIPAHLLDGMGREFWNTNLSLTPVPGOKPNIDVDAVDAQMPAQRIFKEAEKFPVS 300
 Db 259 TGCIPAHLLDGMGREFWNTNLSLTPVPAOKPNIDVDAVDAQMPAQRIFKEAEKFPVS 318
 QY 301 GLPMTGOFWENSMITDPGNVQAVCHPTAMDLGKDFRLMCTKYTMDFLTAHHEGH 360
 Db 319 GLPMTGOFWENSMITDPGNVQAVCHPTAMDLGKDFRLMCTKYTMDFLTAHHEGH 378
 QY 361 IGYMAYRAOPEFLLRNANGFHEAVGEIMSLNATPKHLKSGLLPSDFODESEIINF 420
 Db 379 IGYMAYRAOPEFLLRNANGFHEAVGEIMSLNATPKHLKSGLLPSDFODESEIINF 438
 QY 421 LKQALITVGLPPTMYLEKRWAVFKEGEPKQDMKKWEMKREIVGVEPVPDETC 480
 Db 439 LKQALITVGLPPTMYLEKRWAVFKEGEPKQDMKKWEMKREIVGVEPVPDETC 498
 QY 481 DPASLEHVSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFNML 540
 Db 499 DPASLEHVSNDYSFIRYTRITLYOFQOALCOAKHGEPLHKCDISNSTEAGOKLFNML 558
 QY 541 RIGKSEPTLLENVYGAKNANVRPLNTEPEFTWMLDKONKSEVGMSTWSPYADQSI 600
 Db 559 SIGNSEPTLLENVYGAKNANVRPLNTEPEFTWMLDKONKSEVGMSTWSPYADQSI 618
 QY 601 KVRISLKSALGDKAYEMNDENKYLFRSSVAYAMQOYLKYNQMLLGEEDVAVANLKR 660
 Db 619 KVRISLKSALGDKAYEMNDENKYLFRSSVAYAMQOYLKYNQMLLGEEDVAVANLKR 678
 QY 661 ISFNFPVAPKVNDSIIPRTVEYKARMSRSRINDAFRLNDSLEFLGIOTPLGPPNPQ 720
 Db 679 VSFFYFVAPKVNDSIIPRTVEYKARMSRSRINDAFRLNDSLEFLGIOTPLGPPNPQ 738

RESULT 10

US-10-028-072-72
 ; Sequence 72, Application US/10028072
 ; Publication No. US2003000431A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerltsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Matanabe, Collin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang
 ; TITLE OR INVENTION:
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/028,072
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113

Page 7

1	PRIOR APPLICATION NUMBER: 60/069693
2	PRIOR FILING DATE: 1997-12-16
3	PRIOR APPLICATION NUMBER: 60/072320
4	PRIOR FILING DATE: 1998-01-23
5	PRIOR APPLICATION NUMBER: 60/073612
6	PRIOR FILING DATE: 1998-02-04
7	PRIOR APPLICATION NUMBER: 60/074086
8	PRIOR FILING DATE: 1998-02-09
9	PRIOR APPLICATION NUMBER: 60/074092
10	PRIOR FILING DATE: 1998-02-09
11	PRIOR APPLICATION NUMBER: 60/077791
12	PRIOR FILING DATE: 1998-03-12
13	PRIOR APPLICATION NUMBER: 60/078910
14	PRIOR FILING DATE: 1998-03-20
15	PRIOR APPLICATION NUMBER: 60/079294
16	PRIOR FILING DATE: 1998-03-25
17	PRIOR APPLICATION NUMBER: 60/079663
18	PRIOR FILING DATE: 1998-02-27
19	PRIOR APPLICATION NUMBER: 60/079728
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/080165
22	PRIOR FILING DATE: 1998-03-31
23	PRIOR APPLICATION NUMBER: 60/081203
24	PRIOR FILING DATE: 1998-04-09
25	PRIOR APPLICATION NUMBER: 60/081229
26	PRIOR FILING DATE: 1998-04-09
27	PRIOR APPLICATION NUMBER: 60/081695
28	PRIOR FILING DATE: 1998-04-14
29	PRIOR APPLICATION NUMBER: 60/081817
30	PRIOR FILING DATE: 1998-04-15
31	PRIOR APPLICATION NUMBER: 60/081818
32	PRIOR FILING DATE: 1998-04-15
33	PRIOR APPLICATION NUMBER: 60/082999
34	PRIOR FILING DATE: 1998-04-24
35	PRIOR APPLICATION NUMBER: 60/083322
36	PRIOR FILING DATE: 1998-04-28
37	PRIOR APPLICATION NUMBER: 60/083545
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/084600
40	PRIOR FILING DATE: 1998-05-07
41	PRIOR APPLICATION NUMBER: 60/084627
42	PRIOR FILING DATE: 1998-05-07
43	PRIOR APPLICATION NUMBER: 60/084637
44	PRIOR FILING DATE: 1998-05-07
45	PRIOR APPLICATION NUMBER: 60/085149
46	PRIOR FILING DATE: 1998-05-12
47	PRIOR APPLICATION NUMBER: 60/085323
48	PRIOR FILING DATE: 1998-05-13
49	PRIOR APPLICATION NUMBER: 60/085338
50	PRIOR FILING DATE: 1998-05-13
51	PRIOR APPLICATION NUMBER: 60/085339
52	PRIOR FILING DATE: 1998-05-13
53	PRIOR APPLICATION NUMBER: 60/085579
54	PRIOR FILING DATE: 1998-05-15
55	PRIOR APPLICATION NUMBER: 60/085697
56	PRIOR FILING DATE: 1998-05-15
57	PRIOR APPLICATION NUMBER: 60/085704
58	PRIOR FILING DATE: 1998-05-15
59	PRIOR APPLICATION NUMBER: 60/086414
60	PRIOR FILING DATE: 1998-05-22
61	PRIOR APPLICATION NUMBER: 60/086430
62	PRIOR FILING DATE: 1998-05-22
63	PRIOR APPLICATION NUMBER: 60/087106
64	PRIOR FILING DATE: 1998-05-28
65	PRIOR APPLICATION NUMBER: 60/088026
66	PRIOR FILING DATE: 1998-06-04
67	PRIOR APPLICATION NUMBER: 60/088730
68	PRIOR FILING DATE: 1998-06-10
69	PRIOR APPLICATION NUMBER: 60/088741
70	PRIOR FILING DATE: 1998-06-10
71	PRIOR APPLICATION NUMBER: 60/088810
72	PRIOR FILING DATE: 1998-06-10
73	PRIOR APPLICATION NUMBER: 60/088858

Tue May 27 09:36:31 2003

us-09-978-385-2_copy_19_738.rapb

Page 8

PRIOR FILING DATE: 19/98-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 STIEBOATFLDKFHEAEDLFYOSLASWNTNITEENVONNNAAGKWSAFLEKOST 60
19 STIEBOATFLDKFHEAEDLFYOSLASWNTNITEENVONNNAAGKWSAFLEKOST 78
61 LAQMTPLDIEIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPDNP 120
79 LAQMTPLDIEIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPDNP 138
121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYVLKEMARANYED 180
139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYVLKEMARANYED 198
181 YGDYMRGDEYVNGVDYDYSRGQLIEDVEHTEFEETIKPLYEHLHAYRAKLMNAPSYISP 240
199 YGDYMRGDEYVNGVDYDYSRGQLIEDVEHTEFEETIKPLYEHLHAYRAKLMNAPSYISP 258
241 IGLPRAHLIGDMWGRFTNLVSLVTFGOKRPNIDVTDAVDQAMDQORIFKEAEKFEVSV 300
259 IGLPRAHLIGDMWGRFTNLVSLVTFGOKRPNIDVTDAVDQAMDQORIFKEAEKFEVSV 318
301 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRIIMCTKVYMDDELTAHHEMGH 360
319 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRIIMCTKVYMDDELTAHHEMGH 378
361 IYDMDAYAAOPFLIRNGANGEGHEAVGEIMLSAATPKHLKSIGLSPFOEDNETELNF 420
379 IYDMDAYAAOPFLIRNGANGEGHEAVGEIMLSAATPKHLKSIGLSPFOEDNETELNF 438
421 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREITVGVVEVPDDETYC 480
439 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREITVGVVEVPDDETYC 498
481 DPASLFHVSDYSFIRYRTILYQFOFQALCOAAKHBEPLHKCDISNSTEAGOKL 536
499 DPASLFHVSDYSFIRYRTILYQFOFQALCOAAKHBEPLHKCDISNSTEAGOKL 554

RESULT 11
US-10-121-049-72
Sequence 72, Application US/10121049
Publication No. US2003002239A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Guiney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C17
CURRENT APPLICATION NUMBER: US/10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 STIEBOATFLDKFHEAEDLFYOSLASWNTNITEENVONNNAAGKWSAFLEKOST 60
19 STIEBOATFLDKFHEAEDLFYOSLASWNTNITEENVONNNAAGKWSAFLEKOST 78
61 LAQMTPLDIEIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPDNP 120
79 LAQMTPLDIEIOMLVKLOLQALQONGSSVLESDSKRLNTLNTMSTIYSTGVCNPDNP 138
121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYVLKEMARANYED 180
139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVSKOLRPLYEEVYVLKEMARANYED 198
181 YGDYMRGDEYVNGVDYDYSRGQLIEDVEHTEFEETIKPLYEHLHAYRAKLMNAPSYISP 240
199 YGDYMRGDEYVNGVDYDYSRGQLIEDVEHTEFEETIKPLYEHLHAYRAKLMNAPSYISP 258
241 IGLPRAHLIGDMWGRFTNLVSLVTFGOKRPNIDVTDAVDQAMDQORIFKEAEKFEVSV 300
259 IGLPRAHLIGDMWGRFTNLVSLVTFGOKRPNIDVTDAVDQAMDQORIFKEAEKFEVSV 318
301 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRIIMCTKVYMDDELTAHHEMGH 360
319 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKDFRIIMCTKVYMDDELTAHHEMGH 378
361 IYDMDAYAAOPFLIRNGANGEGHEAVGEIMLSAATPKHLKSIGLSPFOEDNETELNF 420
379 IYDMDAYAAOPFLIRNGANGEGHEAVGEIMLSAATPKHLKSIGLSPFOEDNETELNF 438
421 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREITVGVVEVPDDETYC 480
439 LKQALTIYGTLPFTYMLEKRWMTVFKGELIPKQDMKKWEMKREITVGVVEVPDDETYC 498
481 DPASLFHVSDYSFIRYRTILYQFOFQALCOAAKHBEPLHKCDISNSTEAGOKL 536
499 DPASLFHVSDYSFIRYRTILYQFOFQALCOAAKHBEPLHKCDISNSTEAGOKL 554

RESULT 12
US-10-123-904-72
Sequence 72, Application US/10123904

```

: Publication No. US20030022328A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C54
: CURRENT APPLICATION NUMBER: US/10/123,904
: PRIOR FILING DATE: 2002-04-16
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-10-123-904-72

```

```

Query Match      74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STIEQAKTFLDKFHEADLFYQSSLASWNTNTITEENVONNAGDKSAFLKEQST 60
DB 19 STIEQAKTFLDKFHEADLFYQSSLASWNTNTITEENVONNAGDKSAFLKEQST 78
QY 61 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 120
DB 79 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 138
QY 121 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARAHYED 180
DB 139 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARAHYED 198
QY 181 YGDYMRGDEYVNGVGYYSRGQLEDEYHEFEIKPLYEHLHAYVRKLMNAAPSYISP 240
DB 199 YGDYMRGDEYVNGVGYYSRGQLEDEYHEFEIKPLYEHLHAYVRKLMNAAPSYISP 258
QY 241 IGCPLPHLLGDMGGRFTWNLXSLTFVFGOKPNIDVTDAVDAQDAQRIFEAKKFFVS 300
DB 259 IGCPLPHLLGDMGGRFTWNLXSLTFVFGOKPNIDVTDAVDAQDAQRIFEAKKFFVS 318
QY 301 GLPNMTQGFWNSMLTDPGNVOKAVCHPTAMDIGKDFRILMCTKYTMDFLTAHHBMGH 360
DB 319 GLPNMTQGFWNSMLTDPGNVOKAVCHPTAMDIGKDFRILMCTKYTMDFLTAHHBMGH 378
QY 361 IYDMMAYAAQFPLLRNGANEGFHEAVGIMSLSAATPKHLKISIGLSPDOEDNETEINF 420
DB 379 IYDMMAYAAQFPLLRNGANEGFHEAVGIMSLSAATPKHLKISIGLSPDOEDNETEINF 438
QY 421 LKQALTIYGLTPEFTYMLEKRWMMVFKEIPKQDMKKWMEKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGLTPEFTYMLEKRWMMVFKEIPKQDMKKWMEKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHYSNDYSFIRYTRTLTYOFQFOALCOAAKHGEPILKCDISNSTEAGOKL 536
DB 499 DPASLFHYSNDYSFIRYTRTLTYOFQFOALCOAAKHGEPILKCDISNSTEAGOKL 554

```

```

RESULT 13
US-10-140-470-72

```

```

: Sequence 72, Application US/10140470
: Publication No. US20030022331A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C160
: CURRENT APPLICATION NUMBER: US/10/140,470
: PRIOR FILING DATE: 2002-05-06
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-10-140-470-72

```

```

Query Match      74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STIEQAKTFLDKFHEADLFYQSSLASWNTNTITEENVONNAGDKSAFLKEQST 60
DB 19 STIEQAKTFLDKFHEADLFYQSSLASWNTNTITEENVONNAGDKSAFLKEQST 78
QY 61 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 120
DB 79 LAQMTPLQEIQNTLVKIQLOALQNGSSVLSSEKSKRLNTLNTSTIYSGKCNPNP 138
QY 121 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARAHYED 180
DB 139 QECILLEPGLNEMANSIDYNERLWAMESWSEVQKQRLPYEEYVVLKNEMARAHYED 198
QY 181 YGDYMRGDEYVNGVGYYSRGQLEDEYHEFEIKPLYEHLHAYVRKLMNAAPSYISP 240
DB 199 YGDYMRGDEYVNGVGYYSRGQLEDEYHEFEIKPLYEHLHAYVRKLMNAAPSYISP 258
QY 241 IGCPLPHLLGDMGGRFTWNLXSLTFVFGOKPNIDVTDAVDAQDAQRIFEAKKFFVS 300
DB 259 IGCPLPHLLGDMGGRFTWNLXSLTFVFGOKPNIDVTDAVDAQDAQRIFEAKKFFVS 318
QY 301 GLPNMTQGFWNSMLTDPGNVOKAVCHPTAMDIGKDFRILMCTKYTMDFLTAHHBMGH 360
DB 319 GLPNMTQGFWNSMLTDPGNVOKAVCHPTAMDIGKDFRILMCTKYTMDFLTAHHBMGH 378
QY 361 IYDMMAYAAQFPLLRNGANEGFHEAVGIMSLSAATPKHLKISIGLSPDOEDNETEINF 420
DB 379 IYDMMAYAAQFPLLRNGANEGFHEAVGIMSLSAATPKHLKISIGLSPDOEDNETEINF 438
QY 421 LKQALTIYGLTPEFTYMLEKRWMMVFKEIPKQDMKKWMEKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGLTPEFTYMLEKRWMMVFKEIPKQDMKKWMEKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHYSNDYSFIRYTRTLTYOFQFOALCOAAKHGEPILKCDISNSTEAGOKL 536
DB 499 DPASLFHYSNDYSFIRYTRTLTYOFQFOALCOAAKHGEPILKCDISNSTEAGOKL 554

```

DB 499 DPASLFHVSDDYSFIRYRTTLXOFQFOALCOAKHGEPLHKDISNSTEAGOKL 554

RESULT 14

US-10-175-746-72
Sequence 72, Application US/10175746
Publication No. US20030027270A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-746-72

Query Match

Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 60
DB 19 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 78
QY 61 LAQWYPLQEIOMLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSGKYCNPDP 120
DB 79 LAQWYPLQEIOMLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSGKYCNPDP 138
QY 121 QECILLEPGLINEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 180
DB 139 QECILLEPGLINEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 198
QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLNAAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLNAAVPSYIS 258
QY 241 IGCPLPAHLIGDMGGRFWNTLYSLTVPGOKPNIIDVTDAWVQAMDAORIFKEAEKFPVS 300
DB 259 IGCPLPAHLIGDMGGRFWNTLYSLTVPGOKPNIIDVTDAWVQAMDAORIFKEAEKFPVS 318
QY 301 GLPNNTOGFWENSMILDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 360
DB 319 GLPNNTOGFWENSMILDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 378
QY 361 IQYDMAVAAQPFLLRNGANGEGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNETEINF 420
DB 379 IQYDMAVAAQPFLLRNGANGEGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNETEINF 438
QY 421 LKQALTIYGTLPFTYMLEKRWMMVFEGELIPKDOMMKWMEKREIYGVVPEVPHDETYC 480
DB 439 LKQALTIYGTLPFTYMLEKRWMMVFEGELIPKDOMMKWMEKREIYGVVPEVPHDETYC 498

QY 481 DPASLFHVSDDYSFIRYRTTLXOFQFOALCOAKHGEPLHKDISNSTEAGOKL 536
DB 499 DPASLFHVSDDYSFIRYRTTLXOFQFOALCOAKHGEPLHKDISNSTEAGOKL 554

RESULT 15

US-10-176-918-72
Sequence 72, Application US/10176918
Publication No. US20030027275A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Laureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-918-72

Query Match

Best Local Similarity 99.8%; Pred. No. 5.3e-233;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 60
DB 19 STEEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQNNAGDKMSAFLEKQST 78
QY 61 LAQWYPLQEIOMLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSGKYCNPDP 120
DB 79 LAQWYPLQEIOMLVKQLQALQONGSSVLSSEKSKRLNTILNTMSTIYSGKYCNPDP 138
QY 121 QECILLEPGLINEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 180
DB 139 QECILLEPGLINEIMANSIDYNERLWAMESRSEVQKQRLPYEEYVVLKEMARANYED 198
QY 181 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLNAAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDYDSRGQLIEDVHTFEETIKPLYEHLAAVRAKLNAAVPSYIS 258
QY 241 IGCPLPAHLIGDMGGRFWNTLYSLTVPGOKPNIIDVTDAWVQAMDAORIFKEAEKFPVS 300
DB 259 IGCPLPAHLIGDMGGRFWNTLYSLTVPGOKPNIIDVTDAWVQAMDAORIFKEAEKFPVS 318
QY 301 GLPNNTOGFWENSMILDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 360
DB 319 GLPNNTOGFWENSMILDPGNVQKAVCHPTAWDLGKDFRILMCTKVTYMDDEFLTAHHEMGH 378
QY 361 IQYDMAVAAQPFLLRNGANGEGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNETEINF 420
DB 379 IQYDMAVAAQPFLLRNGANGEGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNETEINF 438

QY 421 LKQALTYGTLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVPEVPHDEYTC 480
DB 439 LKQALTYGTLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVPEVPHDEYTC 498
QY 481 DPASLFHVSNDYSFIRYTRTYLQFOFOBALCOAHKHBGLHKDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYTRTYLQFOFOBALCOAHKHBGLHKDISNSTEAGOKL 554

RESULT 16
US-10-176-921-72

; Sequence 72, Application US/10176921

; Publication No. US20030027276A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Laureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Geriltsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C288

; CURRENT APPLICATION NUMBER: US/10/176,921

; PRIORITY FILING DATE: 2002-06-20

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 72

; LENGTH: 555

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-921-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFELDKFNHEAEDLFYOSLSASWNTNTITEENVOANNAGDKMSAFLKEQST 60
DB 19 STEEQAKTFELDKFNHEAEDLFYOSLSASWNTNTITEENVOANNAGDKMSAFLKEQST 78
QY 61 LAOMYPLQEIONTLVKLOLQALQONGSVLSEDSKRLNTILNTMSTIYSTGKVCNPNP 120
DB 79 LAOMYPLQEIONTLVKLOLQALQONGSVLSEDSKRLNTILNTMSTIYSTGKVCNPNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKOLRPLYEYVVLKNEMARANHYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKOLRPLYEYVVLKNEMARANHYED 198
QY 181 YGDYWRGDEYVNGVDGYDSRGQLIEDVEHTEFEIKPLYEHLHAYVRAKLMAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDGYDSRGQLIEDVEHTEFEIKPLYEHLHAYVRAKLMAVPSYIS 258
QY 241 IGCPLPAHLGDMWGRFTNLTSLVPGOKPNIDVTDAVDAQDAORIFKEAEKFFVS 300
DB 259 IGCPLPAHLGDMWGRFTNLTSLVPGOKPNIDVTDAVDAQDAORIFKEAEKFFVS 318
QY 301 GLPNNTOGFWNSMLTDGNNVOKAVCHPTAMDGLGDFRILMCTKYTMDDFLTAHHEMGH 360
DB 319 GLPNNTOGFWNSMLTDGNNVOKAVCHPTAMDGLGDFRILMCTKYTMDDFLTAHHEMGH 378
QY 361 IYDAMAYAAOPFLIRNGANEGHEAVGEIMSLAATPKHLISGILSPDFEDNETEINF 420

DB 379 IYDAMAYAAOPFLIRNGANEGHEAVGEIMSLAATPKHLISGILSPDFEDNETEINF 438
QY 421 LKQALTYGTLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVPEVPHDEYTC 480
DB 439 LKQALTYGTLPFTYMLEKRMWVFKGEIPKQDMKKWMEKREIVGVPEVPHDEYTC 498
QY 481 DPASLFHVSNDYSFIRYTRTYLQFOFOBALCOAHKHBGLHKDISNSTEAGOKL 536
DB 499 DPASLFHVSNDYSFIRYTRTYLQFOFOBALCOAHKHBGLHKDISNSTEAGOKL 554

RESULT 17
US-10-137-865-72

; Sequence 72, Application US/10137865

; Publication No. US20030032155A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Laureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Geriltsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C154

; CURRENT APPLICATION NUMBER: US/10/137,865

; PRIORITY FILING DATE: 2002-05-03

; PRIOR APPLICATION REMOVED - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 72

; LENGTH: 555

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-137-865-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STEEQAKTFELDKFNHEAEDLFYOSLSASWNTNTITEENVOANNAGDKMSAFLKEQST 60
DB 19 STEEQAKTFELDKFNHEAEDLFYOSLSASWNTNTITEENVOANNAGDKMSAFLKEQST 78
QY 61 LAOMYPLQEIONTLVKLOLQALQONGSVLSEDSKRLNTILNTMSTIYSTGKVCNPNP 120
DB 79 LAOMYPLQEIONTLVKLOLQALQONGSVLSEDSKRLNTILNTMSTIYSTGKVCNPNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKOLRPLYEYVVLKNEMARANHYED 180
DB 139 QECILLEPGLNEIMANSIDYNERLWAMESWSEVGKOLRPLYEYVVLKNEMARANHYED 198
QY 181 YGDYWRGDEYVNGVDGYDSRGQLIEDVEHTEFEIKPLYEHLHAYVRAKLMAVPSYIS 240
DB 199 YGDYWRGDEYVNGVDGYDSRGQLIEDVEHTEFEIKPLYEHLHAYVRAKLMAVPSYIS 258
QY 241 IGCPLPAHLGDMWGRFTNLTSLVPGOKPNIDVTDAVDAQDAORIFKEAEKFFVS 300
DB 259 IGCPLPAHLGDMWGRFTNLTSLVPGOKPNIDVTDAVDAQDAORIFKEAEKFFVS 318
QY 301 GLPNNTOGFWNSMLTDGNNVOKAVCHPTAMDGLGDFRILMCTKYTMDDFLTAHHEMGH 360

```
Db 319 GLPNMTQGFENSMULTDPGNVOKAVCHPTAMDICKGDFRILMCTKYTMDFLTAHHKMGH 378
Qy 361 IQYDMAVAAOPELLRNGANEGFHEAVEIGMSLSAATPKHLKSIIGLSPDFOEDENETEINF 420
Db 379 IQYDMAVAAOPELLRNGANEGFHEAVEIGMSLSAATPKHLKSIIGLSPDFOEDENETEINF 438
Qy 421 LKQALITVGLPTPTMYLMEKRWMPFKGEIPKQOMMKWMEKREIYGVEVPHPDETTC 480
Db 439 LKQALITVGLPTPTMYLMEKRWMPFKGEIPKQOMMKWMEKREIYGVEVPHPDETTC 498
Qy 481 DPASLFHVSNDYSFIRYTRITLYOFQFOEALCOAAKHGRLKCDISNSTEAGOKL 536
Db 499 DPASLFHVSNDYSFIRYTRITLYOFQFOEALCOAAKHGRLKCDISNSTEAGOKL 554

RESULT 18
US-10-474-72
; Sequence 72, Application US/10140474
; Publication No. US20030032156A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey J.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C162
CURRENT APPLICATION NUMBER: US/10/140,474
CURRENT FILING DATE: 2002-05-06
PRIORITY APPLICATION removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-474-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STIEQAKTFLDKNHEADLEFYOSSLSASWNTNTITEENVONMNAAGDKSAFLKEOST 60
Db 19 STIEQAKTFLDKNHEADLEFYOSSLSASWNTNTITEENVONMNAAGDKSAFLKEOST 78
Qy 61 LAQMPLOEIONLTVKIQALQOONGSSVLSSEDKSKRLNTILNTMSTIYSTGKVCNPDP 120
Db 79 LAQMPLOEIONLTVKIQALQOONGSSVLSSEDKSKRLNTILNTMSTIYSTGKVCNPDP 138
Qy 121 QECILLEPGLNEIMANSIDYNERLMAWESRSEVGKOLRPLYEYVYLKNEKARAHYED 180
Db 139 QECILLEPGLNEIMANSIDYNERLMAWESRSEVGKOLRPLYEYVYLKNEKARAHYED 198
Qy 181 YGDWVRGDEYVNGVDYDSRGQLIEDVHEETFEETIKPLYEHLHAYVRAKLMAVPSYIS 240
Db 199 YGDWVRGDEYVNGVDYDSRGQLIEDVHEETFEETIKPLYEHLHAYVRAKLMAVPSYIS 258
Qy 241 IGCPLPAHLGDMGWFNTNLSLTVPGQKPNIDVTAMDVAQDAORIFKEAKFPVSV 300
Db 259 IGCPLPAHLGDMGWFNTNLSLTVPGQKPNIDVTAMDVAQDAORIFKEAKFPVSV 318
```

```
Qy 301 GLPNMTQGFENSMULTDPGNVOKAVCHPTAMDICKGDFRILMCTKYTMDFLTAHHKMGH 360
Db 319 GLPNMTQGFENSMULTDPGNVOKAVCHPTAMDICKGDFRILMCTKYTMDFLTAHHKMGH 378
Qy 361 IQYDMAVAAOPELLRNGANEGFHEAVEIGMSLSAATPKHLKSIIGLSPDFOEDENETEINF 420
Db 379 IQYDMAVAAOPELLRNGANEGFHEAVEIGMSLSAATPKHLKSIIGLSPDFOEDENETEINF 438
Qy 421 LKQALITVGLPTPTMYLMEKRWMPFKGEIPKQOMMKWMEKREIYGVEVPHPDETTC 480
Db 439 LKQALITVGLPTPTMYLMEKRWMPFKGEIPKQOMMKWMEKREIYGVEVPHPDETTC 498
Qy 481 DPASLFHVSNDYSFIRYTRITLYOFQFOEALCOAAKHGRLKCDISNSTEAGOKL 536
Db 499 DPASLFHVSNDYSFIRYTRITLYOFQFOEALCOAAKHGRLKCDISNSTEAGOKL 554
```

```
RESULT 19
US-10-142-431-72
; Sequence 72, Application US/10142431
; Publication No. US20030036179A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey J.
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C251
CURRENT APPLICATION NUMBER: US/10/142,431
CURRENT FILING DATE: 2002-05-10
PRIORITY APPLICATION removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-142-431-72

Query Match 74.9%; Score 2897; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 5.3e-232;
Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STIEQAKTFLDKNHEADLEFYOSSLSASWNTNTITEENVONMNAAGDKSAFLKEOST 60
Db 19 STIEQAKTFLDKNHEADLEFYOSSLSASWNTNTITEENVONMNAAGDKSAFLKEOST 78
Qy 61 LAQMPLOEIONLTVKIQALQOONGSSVLSSEDKSKRLNTILNTMSTIYSTGKVCNPDP 120
Db 79 LAQMPLOEIONLTVKIQALQOONGSSVLSSEDKSKRLNTILNTMSTIYSTGKVCNPDP 138
Qy 121 QECILLEPGLNEIMANSIDYNERLMAWESRSEVGKOLRPLYEYVYLKNEKARAHYED 180
Db 139 QECILLEPGLNEIMANSIDYNERLMAWESRSEVGKOLRPLYEYVYLKNEKARAHYED 198
Qy 181 YGDWVRGDEYVNGVDYDSRGQLIEDVHEETFEETIKPLYEHLHAYVRAKLMAVPSYIS 240
Db 199 YGDWVRGDEYVNGVDYDSRGQLIEDVHEETFEETIKPLYEHLHAYVRAKLMAVPSYIS 258
```



```

QY 241 IGCIPAHLLGDMGREFWTLNLSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 300
DB 259 IGCIPAHLLGDMGREFWTLNLSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 318
QY 301 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGFRLIMCTKYTMDDFLTAHHMGH 360
DB 319 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGFRLIMCTKYTMDDFLTAHHMGH 378
QY 361 IGYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNTEINF 420
DB 379 IGYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNTEINF 438
QY 421 LKQALTIYGTLPETYMLEKRMWVFKGEIPKQDMKKWMEKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGTLPETYMLEKRMWVFKGEIPKQDMKKWMEKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHVSDYSFIRYRTLYQFOFOEALCOAAKHGFLHKCDISNSTEAGOKL 536
DB 499 DPASLFHVSDYSFIRYRTLYQFOFOEALCOAAKHGFLHKCDISNSTEAGOKL 554

```

RESULT 20

```

US-10-143-114-72
; Sequence 72, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-114-72

```

Query Match 74.9%; Score 2897; DB 9; Length 555;
 Best Local Similarity 99.8%; Pred. No. 5.3e-232;

Matches 535; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 STIEBAKFTLDKFNHEADLFYQSSLASWNTNTITEENVONMANNAGDKMSAFLEKQST 60
DB 19 STIEBAKFTLDKFNHEADLFYQSSLASWNTNTITEENVONMANNAGDKMSAFLEKQST 78
QY 61 LAQWTFLOETONLTVKLOLQALQONGSSVLSSEKSKRLNTINTMSTIYSTGKVCNPDP 120
DB 79 LAQWTFLOETONLTVKLOLQALQONGSSVLSSEKSKRLNTINTMSTIYSTGKVCNPDP 138
QY 121 QECILLEPGLNEMANSIDYNERLMAWMSRSEVGKQRLPLYEETVVLKNEMARAHYED 180
DB 139 QECILLEPGLNEMANSIDYNERLMAWMSRSEVGKQRLPLYEETVVLKNEMARAHYED 198
QY 181 YGDYWRGDEYVNGVDYDYSRQGLIEDVEHTEFEIKPLYEHLHAVVRAKLMNAVPSYISP 240

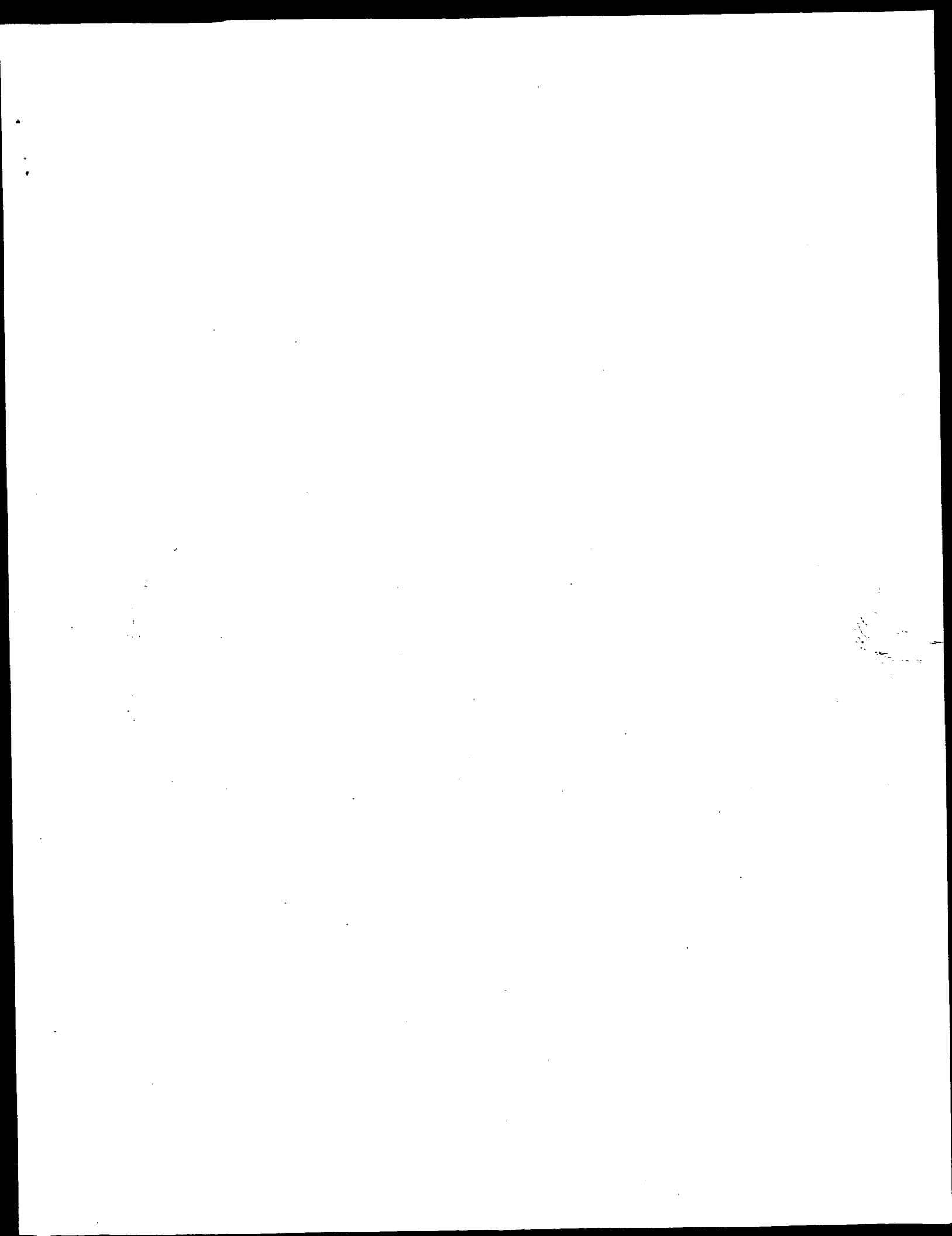
```

```

DB 199 YGDYWRGDEYVNGVDYDYSRQGLIEDVEHTEFEIKPLYEHLHAVVRAKLMNAVPSYISP 258
QY 241 IGCIPAHLLGDMGREFWTLNLSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 300
DB 259 IGCIPAHLLGDMGREFWTLNLSLVPEFGOKPNIDVTAMVDQAMDARIFKEAEKFEVS 318
QY 301 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGFRLIMCTKYTMDDFLTAHHMGH 360
DB 319 GLPNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGFRLIMCTKYTMDDFLTAHHMGH 378
QY 361 IGYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNTEINF 420
DB 379 IGYDMAYAAOPFLLRNGANGFHEAVGEIMSLSAATPKHLKSGILSPDFOEDNTEINF 438
QY 421 LKQALTIYGTLPETYMLEKRMWVFKGEIPKQDMKKWMEKREIYGVVEPVPHDETYC 480
DB 439 LKQALTIYGTLPETYMLEKRMWVFKGEIPKQDMKKWMEKREIYGVVEPVPHDETYC 498
QY 481 DPASLFHVSDYSFIRYRTLYQFOFOEALCOAAKHGFLHKCDISNSTEAGOKL 536
DB 499 DPASLFHVSDYSFIRYRTLYQFOFOEALCOAAKHGFLHKCDISNSTEAGOKL 554

```

Search completed: May 26, 2003, 17:59:11
 Job time : 41 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 17:42:18 ; Search time 37 Seconds

(without alignments)
1870.724 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 1 STTBQAKTFLDKFHEAD.....DNSLEPLGIQPTLGPPNPP 720

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:73:*
2: PIR:*
3: PIR:*
4: PIR:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3869	100.0	804	2	T14762
2	1337	34.6	732	1	S05238
3	1337	34.6	1306	1	A31759
4	1334	34.5	732	1	A35655
5	1334	34.5	1312	1	A34171
6	1312	33.9	1193	2	JC2489
7	1310	33.9	1313	1	JC2038
8	1283	33.2	737	1	A34402
9	1283	33.2	1309	1	S35484
10	1054.5	27.3	611	2	S65472
11	1027.5	26.6	630	2	JC5374
12	1022	26.4	615	2	A57533
13	638.5	16.5	907	2	T15792
14	157	4.1	532	2	C83696
15	154	4.0	502	2	AF1310
16	147	3.8	502	2	AE1682
17	139.5	3.6	987	2	AI2011
18	139	3.6	987	2	B82938
19	136	3.5	611	2	D82881
20	135	3.5	611	2	D69943
21	125	3.2	627	1	S40048
22	124.5	3.2	987	2	I48373
23	124	3.2	538	2	E72561
24	122	3.2	642	2	E98000
25	121	3.1	607	2	AB3511
26	121	3.1	1034	2	T30574
27	119	3.0	990	2	S23416
28	118	3.0	642	2	G95129
29	117.5	3.0	1339	2	A84683

Hit No. 1/5
too late

30	116	3.0	963	2	C90535	conserved hypothet
31	115.5	3.0	611	2	A75573	probable oligomero
32	115.5	3.0	3433	1	S28381	utrophin - human
33	115	3.0	950	2	A71855	hypothetical prote
34	115	3.0	1780	2	T17272	hypothetical prote
35	114.5	2.9	3856	2	T51174	ataxia-telangiecta
36	114	2.9	685	2	F75370	oligoendopeptidase
37	113.5	2.9	3655	2	T38084	TRAP-like protein
38	113	2.9	772	2	A10968	probable glycosyl
39	113	2.9	902	2	E90270	conserved hypothet
40	113	2.9	1575	2	G82905	conserved hypothet
41	113	2.9	6669	2	S55024	nebulin, skeletal
42	112.5	2.9	808	2	A82877	hypothetical prote
43	112	2.9	524	2	B82202	thermostable carbo
44	111.5	2.9	282	2	H97226	protein containing
45	111	2.9	515	2	C71158	probable thermosta

ALIGNMENTS

RESULT 1

T14762
hypothetical protein DKFp434A014.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T14762

R/Wambuit, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A/Accession: T14762

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-804 <RAW>

A/Cross-references: EMBL:AL110224

A/Experimental source: adult testis; clone DKFp434A014

C/Genetics:

A/Note: DKFp434A014.1

Query Match	100.0%	Score 3869	DB 2	Length 804
Best Local Similarity	100.0%	Pred. No. 3.5e-265		
Matches 720	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	STTBQAKTFLDKFHEADLFYQSSLASWNTNITEENYQNMNAGDKMSAFLEKST	60	
DB	18	STTBQAKTFLDKFHEADLFYQSSLASWNTNITEENYQNMNAGDKMSAFLEKST	77	
QY	61	LAQMTPLQETQNLTKYKLOALQONGSVLSSEKSKRLNTILNTSTYTGKYNPDNP	120	
DB	78	LAQMTPLQETQNLTKYKLOALQONGSVLSSEKSKRLNTILNTSTYTGKYNPDNP	137	
QY	121	QECLEELGELNITMANSIDYNERLWAMESWSEVQKRLPYEEVYLKKNMAANHEED	180	
DB	138	QECLEELGELNITMANSIDYNERLWAMESWSEVQKRLPYEEVYLKKNMAANHEED	197	
QY	181	YGDYWRGDEYVNGVDYDSRGQLTEDEYHTEFEETKPLYEHLHAYVRKLNANAYSTSP	240	
DB	198	YGDYWRGDEYVNGVDYDSRGQLTEDEYHTEFEETKPLYEHLHAYVRKLNANAYSTSP	257	
QY	241	IGCLPAHLIGDMWGFNTNLSLYVPGQKPNIDVTAMDAMDADQRIFEAKRFVSV	300	
DB	258	IGCLPAHLIGDMWGFNTNLSLYVPGQKPNIDVTAMDAMDADQRIFEAKRFVSV	317	
QY	301	GLPNTQGFWEKSMITDPGNOKAVCHPTAMDIGGDRILMCTKVTMDDELTAHHHGH	360	
DB	318	GLPNTQGFWEKSMITDPGNOKAVCHPTAMDIGGDRILMCTKVTMDDELTAHHHGH	377	
QY	361	IQYDVAAYAAPPLLRNGANGEGFHEAVGEIMSLSATPKHLKSIGLSPDEQEDNETINF	420	
DB	378	IQYDVAAYAAPPLLRNGANGEGFHEAVGEIMSLSATPKHLKSIGLSPDEQEDNETINF	437	
QY	421	LLKQALITVGLPFLPYHLEKRWVFKETIRKQDMKRWMEKREIVGVEPVDHETTC	480	

Db 438 LKQALITVGTLPFTYMLEKRWVFKGEIPKQDMKKWMEKREIYGVVEPYHDETYC 497
 QY 481 DPASLFHNSYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTPAGOKLNNML 540
 Db 498 DPASLFHNSYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTPAGOKLNNML 557
 QY 541 RLKSEPTWTLLENVYGAKNMNVPLNTYEPFTWLDKDNKNSFVGMSTDMSPYADQSI 600
 Db 558 RLKSEPTWTLLENVYGAKNMNVPLNTYEPFTWLDKDNKNSFVGMSTDMSPYADQSI 617
 QY 601 KVRISLKSALDDKAYEMNDNEMYLFRSSVAAMQYFLKYNQNMILGEEVRYANLKKR 660
 Db 618 KVRISLKSALDDKAYEMNDNEMYLFRSSVAAMQYFLKYNQNMILGEEVRYANLKKR 677
 QY 661 ISFNFTAPKKNVSDIIPTEVEKAIKRSRIRIDAFELNDNSLEFLGIQPTLGPNNPP 720
 Db 678 ISFNFTAPKKNVSDIIPTEVEKAIKRSRIRIDAFELNDNSLEFLGIQPTLGPNNPP 737

RESULT 2

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, testicular splice form - human
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 02-Jul-1998 #text_change 18-Jun-1999
 C:Accession: S05238; A33979
 R:Latton, A.L.; Soubrier, F.; Allegrini, J.; Hubert, C.; Corvol, P.; Alhenc-Gelas, F.
 FEBS Lett. 252, 99-104, 1989
 A:Title: The testicular transcript of the angiotensin I-converting enzyme encodes for the
 A:Reference number: S05238; MUID:89338720; PMID:2547653
 A:Accession: S05238
 A:Molecule type: mRNA
 A:Residues: 1-732 <LAT>
 A:Cross-references: EMBL:X16295; NID:928264; PIDN:CAA34362.1; PID:928265
 R:Ehlers, M.R.W.; Fox, E.A.; Strydom, D.J.; Riordan, J.F.
 Proc. Natl. Acad. Sci. U.S.A. 86, 7741-7745, 1989
 A:Title: Molecular cloning of human testicular angiotensin-converting enzyme: the testis
 A:Reference number: A33979; MUID:90046671; PMID:2534286
 A:Accession: A33979
 A:Molecule type: mRNA
 A:Residues: 1-732 <EHL>
 A:Cross-references: GB:M26657; NID:9338666; PIDN:AAA60611.1; PID:9338667
 A:Experimental source: clones R1.2 and T88
 A:Note: neither the complete nucleic acid sequence nor the complete translation are shown
 C:Comment: For the renal and pulmonary splice form, see PIR:A31759.
 C:Genetics:
 A:Gene: GDB:DCPI; ACE
 A:Cross-references: GDB:119840; OMIM:106180
 A:Map position: 17q23-17q23
 C:Function:
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; glycoprotein; metalloproteinase; peptidyl-dipeptide hydrolase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-732/Product: peptidyl dipeptidase I #status predicted <MAT>
 F:686-702/Domain: transmembrane #status predicted <RM>
 F:103,121,140,186,368,617,651/Binding site: carbonyl group (Asn) (covalent) #status predicted
 F:414,418,434/Binding site: zinc, catalytic (His, His, Glu) #status predicted
 F:415/Active site: Glu #status predicted

Query Match 34.6% Score 1337; DB 1; Length 732;
 Best Local Similarity 41.7% Pred. No.2.4e+86;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 2 TIEBAQKFLDKFNHAEADLFYOSLSAMNTNTITE-----NYQNNNMGDKWSA 53
 Db 70 TDEAASKFVEYEDTDSQVWVMEYEAAMNTNTITESTILKQNNHNT----- 123
 QY 54 FLKEGSLAQMTPLOEIONLVKIQLOALQONGSSVLESDSKRLNTILNMTSTYSGK 113
 Db 124 --LKYGTOARKEDVQALQNTYIKRIKIKVQDLERALPAQGLEENKILADMETTYSVAT 181
 QY 114 VCNPNQPCCLLEELGELIMANSLDYNERLWAMSNRSEYKQRLPYEEYVVLKNNMA 173

Db 182 VCHPNQ--SCIDLEPDLTVMMATSRKYEDELWAMEBWRKAGARALLOEPKYEVEILNOAA 239
 QY 174 RANHYEDGDMYRGDVEANGVGYDSRQGLLEDVEHEFEELKPYEHLHAAYRAKLNA 233
 Db 240 RLNGYDAGDSMSMTEPSLE-----QDERLEQELPLNLHAAYRALNRH 289
 QY 234 Y-PSYISPIGCLPAHLLDGMGRFWNTLSLYVPEGQKRNIDVTAMVQDAQRIKRE 292
 Db 290 YGAQHNLNLEGPYIAHLILGNMMAQTNWISYDLVVPFSAQSMOTTEAMLKQGTPRMRE 349
 QY 293 AEKFFVSVGLPNKQGFENSMITDGCNVQKAVCHETANDLQK-DERLTKTKYTMDF 351
 Db 350 ADEFTSLGLPVPPEFVWKSMLKPTDREYVCHASADFYNGDFRDKOCTVYNLDEL 409
 QY 352 LTAHHEGHIQYDMAFAAQPELLRNGANGFPAVEGISTSAAPKHLKSLGSLDPQ 411
 Db 410 VVAHHEGHIQYMQKDEPVALREANGFHEATDVALSVTPKHLISNLISSEGG 469
 QY 412 EDNETEINFLKQALITVGTLPFTYMLEKRWVFKGEIPKQDMKKWMEKREIYGVVE 471
 Db 470 SD-EHIDINFLMKALDKIAFIPSYLVDMQRMRFVFGSITKENYQNMWMLRLKYQGLCP 528
 QY 472 PVHDETYCDPASFLEHNSYSPFRYRTPLTQFOFQALCOAKHGGPLHKCDISNSTE 531
 Db 529 PVRTGDEPDEPKAFHIPSSVPIRYFVSFTIOFQHEALCOAGHTGPLHKCDIYQSK 588
 QY 532 AGQRLNMLRLKSEPTWTLLENVYGAKNMNVPLNTYEPFTWLDKDNK--NSFVGM- 588
 Db 589 AGQRLNMLRLKSEPTWTLLENVYGAKNMNVPLNTYEPFTWLDKDNK--NSFVGM- 588
 QY 589 STDMSPYADQSI 599
 Db 649 QYMWTPNSARS 659

RESULT 3

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, renal and pulmonary splice form - human
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase
 C:Species: Homo sapiens (man)
 C:Date: 07-Jun-1990 #sequence_revision 02-Jul-1998 #text_change 18-Jun-1999
 C:Accession: A31759; P00004
 R:Soubrier, F.; Alhenc-Gelas, F.; Hubert, C.; Allegrini, J.; John, M.; Tregear, G.; C.
 Proc. Natl. Acad. Sci. U.S.A. 85, 9386-9390, 1988
 A:Title: Two putative active centers in human angiotensin I-converting enzyme revealed
 A:Reference number: A31759; MUID:89071703; PMID:2849100
 A:Accession: A31759
 A:Molecule type: mRNA
 A:Residues: 1-1306 <SD>
 A:Cross-references: GB:U04144; NID:9178285; PIDN:AAA51684.1; PID:9178286
 A:Experimental source: kidney
 A:Note: parts of this sequence, including the amino end of the mature protein, were derived
 R:Takeuchi, K.; Shimizu, T.; Ohishi, N.; Seyama, Y.; Takaku, F.; Yotsumoto, H.
 J. Biochem. 106, 442-445, 1989
 A:Title: Purification of human lung angiotensin-converting enzyme by high-performance
 A:Reference number: P00004; MUID:90110025; PMID:2558109
 A:Accession: P00004
 A:Molecule type: protein
 A:Residues: 'XX', 32-34, 'E', 36-37, 'X', 39-41, 'R', 43-46 <TAK>
 A:Experimental source: lung
 C:Comment: This splice form is found in many tissues, in particular kidney and lung v
 C:Genetics:
 A:Gene: GDB:DCPI; ACE
 A:Cross-references: GDB:119840; OMIM:106180
 A:Map position: 17q23-17q23
 C:Function:
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; blood pressure control; glycoprotein; kidney; lung;
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-1306/Product: peptidyl dipeptidase I #status predicted <MAT>

F:1260-1276/Domain: transmembrane #status predicted <TRM>
 F:38,54,74,111,146,160,318,445,509,523,677,695,714,760,942,1191,1225/Binding site: carbo
 F:390,394/Binding site: zinc (His) #status predicted
 F:988,992,1008/Binding site: zinc catalytic (His, His, Glu) #status predicted
 F:989/Active site: Glu #status predicted

Query Match 34.6% Score 1337; DB 1; Length 1306;
 Best Local Similarity 41.7% Pred. No. 5,9e-86;
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

2 TIEQAKFTLDKFNHEADLFYQSSLSAWNTNTEENYNNMAGDKNSA 53
 644 TTEAEASKFEVEEDRTSYVWNEEYAEANNNTNTEENSKILLQKNQIANHT----- 697
 54 FLKEGSTLAQMYPLOEIOMLYKLOALQONGSSVLSDEKSKRLNTLNTSTYSGK 113
 698 --LKGTQAKKFDVNOLOMTTKRIKKVQDERALPAQELKEEYNLILDEETYSVAT 755
 114 VCNPNPQCLLEPGLNEIMANSLDYNERLWAMESWSEVQKOLRPLYEEVYLKNEMA 173
 756 VCHPNG--SCLOLEDPLTVVMAISKRKEDILMAWGMKACRAILQYFVKVELINDQA 813
 174 RANHEEDYGVYRGDTEVANGVDGYDSKGLIEDVEHTEFEIKPLYLHAYVRAKLMNA 233
 814 RLNGYVDAGDSWRSMTETPSLE-----QDLERLFOELQPLYLHAYVRAKLRH 863
 234 Y-PTSTIPGCLPLGLDMGREFWNTLSLTPPGQKPNIDVYAMVQAMDAQRITKE 292
 864 YGAQHINLEGPPIPAHLGNMMAQWTSNITDVAAPPASAPNDITAMIKOGTPRRIRKE 923
 293 AKEFFVSVGLPMTQGFENSMITDPGNVQKAVCHPTAMDGK--DEFILMCTKTYMD 351
 924 ADDEFTSLGLPVPPEFNKSMLEKPTDQREVYCHASAMDPYNGDFIKCTTYVNL 983
 352 LTAHEMHIOYDAAVAAQPELLRNGANGHEAVGELMSLSAATPKHLKISGLSPDFQ 411
 984 VVAHEMHIOYFQYKDLPAALREGANPGEHAIQDIALSVSPKHLISLNTLSRG 1043
 412 EDNEIEINFLKQALTIYGLTFEYMLKRWVYKGEIPKDDMMKKWKKRELIVGYE 471
 1044 SD-EHDINFLMKMALDKIAFIPTSLVDQWRRPFDGISTYENTNQEWMWSRLKYQICP 1102
 472 PVPHEDEYCDPASLFHNSNDYSFIRYTRTYLQFOFQALCOAAKHGEPHKKDISNTE 531
 1103 PVPRIQGDPDGAFHPISSVPIRYEVSFTIQFQHEALCOAGHGPPLKKCOIYQSK 1162
 532 AGOKLENNLRGKSPWTLAENYVGAKNMVRPLNTEFLTWLKDOK--NSFVGW- 588
 1163 AGORLATAMKLGFSRPMPEAMQITGQPNMSASAMLSYFKPLDMLTENELHGEKLGWP 1222
 589 STDMSPYADQS 599
 1223 QYNWTPNSARS 1233

RESULT 4

A35655
 peptidyl-dipeptidase A (EC 3.4.15.1), testis - mouse
 N:Alternate names: peptidyl-dipeptidase I, testis
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 R:Howard, T.E.; Shai, S.Y.; Langford, K.G.; Martin, B.M.; Bernstein, K.E.
 M01. Cell. Biol. 10, 4294-4302, 1990
 A:Title: Transcription of testicular angiotensin-converting enzyme (ACE) is initiated w
 A:Reference number: A35655; M01D:90318396; PMID:2164656
 A:Accession: A35655
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-732 <HOW>
 A:Cross-references: GB:M55333; NID:g191589; PIDN:AAA37149.1; PID:g191590
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein; z

Query Match 34.5% Score 1334; DB 1; Length 732;
 Best Local Similarity 42.6% Pred. No. 3.9e-86;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

2 TIEQAKFTLDKFNHEADLFYQSSLSAWNTNTEENYNNMAGDKNSAFLKEOSTL 61
 69 TDEAKDAFVEEYDRAQVLLNVEYAEANNQNTNTEEGSILLKESTVSNHLKGT 128
 62 AQMYDLOEIOMLYKLOALQONGSSVLSDEKSKRLNTLNTSTYSGKVCNPNPQ 121
 129 AKTDVSNQSSSKRIKILQNDRAVLPKELEENQIILDEETYSLSNICYTG-- 166
 122 ECLLEPGLNEIMANSLDYNERLWAMESWSEVQKOLRPLYEEVYLKNEMARNHEDY 181
 187 TCMPLERPLTVMMAISKRKEDILMAWGMKACRAILPPEPKVFSNNIAKNGYDA 246
 182 GDYRGDYEYVANGVDGYDSKGLIEDVEHTEFEIKPLYLHAYVRAKLMNAVPS-YISP 240
 247 GDSWRSIYESDNL-----QDEKLYQELQPLYLHAYVRSLSRHHGSEYINL 296
 241 IGCPLAHLGLDMGREFWNTLSLTPPGQKPNIDVYAMVQAMDAQRITKEAKFEFVS 300
 297 DGPPIPAHLGNMMAQWTSNITDVAAPPASAPNDITAMIKOGTPRRIRKEADNFTSL 356
 301 GLPMTQGFENSMITDPGNVQKAVCHPTAMDGK--DEFILMCTKTYMDPLTAHEMG 359
 357 GLPVPPEFNKSMLEKPTDQREVYCHASAMDPYNGDFIKCTVSNMDELVAHHEMG 416
 360 HIQYDAAVAAQPELLRNGANGHEAVGELMSLSAATPKHLKISGLSPDFQDNEIEIN 419
 417 HIQFYQYKDLPYFRBAGNPGHEAIGDIALSVSPKHLISLNTLSRGSGYEYDIN 475
 420 FLKQALTIYGLTFEYMLKRWVYKGEIPKDDMMKKWKKRELIVGYEVEVDEY 479
 476 FLMKMALDKIAFIPTSLIDQWRRPFDGISTYENTNQEWMWSRLKYQICPVPSSQD 535
 480 CDPASLFHNSNDYSFIRYTRTYLQFOFQALCOAAKHGEPHKKDISNTEAGKILFVN 539
 536 FDSGSKFHVANPYVRYEVSFTIQFQHEALCRAAGHGPPLKKCOIYQSKRGKILDA 595
 540 LRLKSEPWTLAENYVGAKNMVRPLNTEFLTWLKDOK--NSFVGW-STDMS 594
 596 MKLGSKPPEAMKLLTGQPNMSASAMNIFKPLTEMLVYENRHGETIGWPEYNAP 653

RESULT 5

A34171
 peptidyl-dipeptidase A (EC 3.4.15.1) precursor - mouse
 N:Alternate names: ACE; angiotensin-converting enzyme; carboxypeptidase; dipeptidyl c
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A34171; A29220; A61477
 R:Bernstein, K.E.; Martin, B.M.; Edwards, A.S.; Bernstein, E.A.
 J. Biol. Chem. 264, 11945-11951, 1989
 A:Title: Mouse angiotensin-converting enzyme is a protein composed of two homologous
 A:Reference number: A34171; M01D:89308599; PMID:2545691
 A:Accession: A34171
 A:Molecule type: mRNA
 A:Residues: 1-1312 <BER>
 A:Cross-references: GB:J04947
 R:Bernstein, K.E.; Martin, B.M.; Bernstein, E.A.; Linton, J.; Striker, L.; Striker, G
 J. Biol. Chem. 263, 11021-11024, 1988
 A:Title: The isolation of angiotensin-converting enzyme cDNA.
 A:Reference number: A29220; M01D:88298730; PMID:2841312
 A:Accession: A29220
 A:Molecule type: mRNA
 A:Residues: 1-332 <BE2>
 A:Cross-references: GB:J03940; NID:g191583; PIDN:AAA37146.1; PID:g191584
 R:Bernstein, K.E.; Martin, B.M.; Striker, L.; Striker, G.
 Kidney Int. 33, 652-655, 1988
 A:Title: Partial protein sequence of mouse and bovine kidney angiotensin converting e
 A:Reference number: A61477; M01D:88215372; PMID:2835538

A:Accession: A61477
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 35-54 <BB3>
 A:Experimental source: kidney
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; blood pressure control; membrane protein; peptidyl-dipeptidase I #status predicted <S>
 F:1-34/Domain: signal sequence #status predicted <S>
 F:35-1312/Product: peptidyl dipeptidase I #status predicted <MA>

Query Match 34.5%; Score 1334; DB 1; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 9.6e-86;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 2 TTEQAKTFIDKHEADLEFYQSSLSMNNTNTEENVOGNNAGDKSAFLKEOSTL 61
 DB 649 TDEKADRFVEYRTOVLNEYAEANMOQNTITTEGSKITLKESTEVSNHLKCTR 708
 QY 62 AQMTPLAEIQNTLVKLOLALQNGSSVLSSEKSKRLNTLNTSTYTGKVCNPDPQ 121
 DB 709 AKTIDVSNFQSSIKRIKQNDRAVLPEKLEEVNQLIDMETTYLSNICYTG-- 766
 QY 122 ECLLEPGELNEMSLDYNERLAMESWSEVQKOLRPLEEYVVLKNEKARAHEDY 181
 DB 767 TCMLEPDLTMMATSRKYEELLMAMKSWDRKVRALPEPKYVEFSNKIAKLNGYTD 826
 QY 182 GDYMGDEYVNGVDGYDSRGQLIEDVEHTEEEKIPLEYHLHAYRAKLMAVPS-YISP 240
 DB 827 GDSWRSLYSEDNLE-----QDLKLYQOEIPLYLNLHAYRSLHNGSEYINL 876
 QY 241 IGCLEPAHLIGDMWGFMTNLSLVPEFGKPNIDVTAMVQADAOIRFEAKRFVSV 300
 DB 877 DGPFAHLIGNMMAQSWNIDYLVAFPPSAPINIDATEAMIKQWTPRIEFKEANPFTSL 936
 QY 301 GLPNTQGFEMSMULTDGNVOKACHPDANDLGG-DEFTLMCKVTMDPFLAHNEMG 359
 DB 937 GLPPEPFEMKSMLEKPTDGEVYVCHSANDFYNGKDFRIKQCTSVAMDELVAHHNEMG 996
 QY 360 HIOYDMAYAAQPFLLRNGANGFHEAVGEIEMSLATPRKHLKSGLLSPDFQEDNETEIN 419
 DB 997 HIOYDMQKDLPEVTFREGANGFHEAIGDIMALSVTPKHLYSJNLSTE-GSGEYEDIN 1055
 QY 420 FLTKALITVGLPFTYMLEKRMVVFGEIPEKDOMKMKWEMKREIYGVVEPYPHDET 479
 DB 1056 FLTKALDKIAFIPSYLLIDQRMKRVFEGSTKTEYNQEMWSLKRKTGCLPPRPSGCD 1115
 QY 480 CDPASLFIHNSDYSFIRYTRTLXQFOFQALCQAAREGPHKCDISNSTEAGOKLFNM 539
 DB 1116 PDPSKGFHPANVPYRVFVSIIQFQHEALCRAAGTGPHKCDIYQSKAEAGRLADA 1175
 QY 540 LRLGSEPTWTLLENVGAKNMVRPLNTEPELFTWLKDONK--NSFVGM-STDMSP 594
 DB 1176 KRLGSKWPPEAMKLTICQPNMSASAMMYEKLTEWLVTENRRHGETLIGPEYVNM 1233

RESULT 6
 JC2489
 peptidyl-dipeptidase A (EC 3.4.15.1) - chicken
 N:Alternate names: angiotensin converting enzyme
 C:Species: Gallus gallus (chicken)
 C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 28-May-1999
 C:Accession: JC2489

A:Residues: 1-1193 <EST>
 A:Cross-references: GB:U40175; NID:9685168; PIDN:AAA75554.1; PID:9994708
 C:Comment: This enzyme is a zinc dependant dipeptidyl carboxypeptidase that cleaves a
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: metal binding; peptidyl-dipeptide hydrolase; zinc

F:316,331,914,929/Binding site: zinc, catalytic (Glu, His, Glu, His) #status predicted
 Query Match 33.9%; Score 1312; DB 2; Length 1193;
 Best Local Similarity 40.4%; Pred. No. 3e-84;
 Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

QY 4 EEOAKTFIDKHEADLEFYQSSLSMNNTNTEENVOGNNAGDKSAFLKEOSTLAQ 63
 DB 544 EAOAKFEFLSEYNSTAEVNNAYTEASWEYNTNITDHNKREVMLEKMLASKHTIEGMBAR 603
 QY 64 MYPLQEIQNTLVKLOLALQNGSSVLSSEKSKRLNTLNTSTYTGKVCNPDP--P 120
 DB 604 QEPSPFOEYTRILNKLSVLERALPDELEKETVLTLSDETTYSAKVCRENNTNHP 663
 QY 121 QECLEPGELNEMSLDYNERLAMESWSEVQKOLRPLEEYVVLKNEKARAHEDY 180
 DB 664 -----LDPLDTLILASROYNELLFMKGMWMAASGAKIKDKRYVELSNKAALVNGYTD 718
 QY 181 YGDMRGDEYVNGVDGYDSRGQLIEDVEHTEEEKIPLEYHLHAYRAKLMAV-PSTYS 239
 DB 719 NGAYMSLYETPFEE-----EDLERLYLOIPLYLNLHAYRRLAYKNGAENHS 768
 QY 240 PIGCLPAHLIGDMWGFMTNLSLVPEFGKPNIDVTAMVQADAOIRFEAKRFVSV 299
 DB 769 LKGPFAHLIGNMMAQSWNIDYLVAFPPDQATKVDATPKMKQGWTPKMEFEESDRFTS 828
 QY 300 VGLPNTQGFEMSMULTDGNVOKACHPDANDL-GKGFRLMCKTVMDDPFLAHNEM 358
 DB 829 LGLIPDQEFWDMKSMLEKPTDGEVYVCHSANDFYNGKDFRIKQCTSVAMDELVAHHNEM 888
 QY 359 GHIOYDMAYAAQPFLLRNGANGFHEAVGEIEMSLATPRKHLKSGLLSPDFQEDNETE 418
 DB 889 GHVOYPLQYMDQISPRDGNANPEFHEAIGDIMALSVTPKHLHSLNLD-QYTEDESDI 947
 QY 419 NFLKQALITVGLPFTYMLEKRMVVFGEIPEKDOMKMKWEMKREIYGVVEPYPHDET 478
 DB 948 NLYMSIALDKIAFIPSYLLIDQRMKRVFEGSTKTEYNQEMWSLKRKTGCLPPRPSBED 1007
 QY 479 YCPASLFIHNSDYSFIRYTRTLXQFOFQALCQAAREGPHKCDISNSTEAGOKLFNM 538
 DB 1008 DFPDGAFFHLPANVPYRVFVSIIQFQHEALCRAAGTGPHKCDIYQSKAEAGRLADA 1067
 QY 539 MLRLGSEPTWTLLENVGAKNMVRPLNTEPELFTWL-KDONKNSFVGM-STDMSPY 595
 DB 1068 AMKLGFSKPPEAMKLTICQPNMSASAMMYEKLTEWLVTENRRHGETLIGPEYVNM 1127
 QY 596 ADOSIKYRISLKSALG-----DKAYENDNEMTYLFRSSVAYAMQRYFLAKY 641
 DB 1128 AVTEFHAAITDADFLGWSVGTQKQATAGAW-----VLLALALVFLITISIFLYAK 1175

RESULT 7
 JC2038
 peptidyl-dipeptidase A (EC 3.4.15.1) - rat
 N:Alternate names: angiotensin converting enzyme; kininase II
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: JC2038

A:Residues: 1-1313 <KOI>
 A:Cross-references: GB:U03734; NID:9437289; PIDN:AAA8211.1; PID:9437290
 A:Note: the authors translated the codon AGC for residue 159 as Tyr
 C:Comment: This enzyme is a zinc-containing dipeptidyl carboxypeptidase that cleaves angiotensin
 C:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein;
 F:393-400,990-998/Region: catalytic #status predicted
 F:1264-1284/Domain: transmembrane #status predicted <TM>

Query Match 33.9%; Score 1310; DB 1; Length 1313;
 Best Local Similarity 42.0%; Pred. No. 4.8e-84;
 Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

2 TIEBOAKTFLEKDNHEADLFYQSSLSAMNTNTTEENVOMNNAQKMAFLKEOSTL 61
 DB TDEAKANFVEEYEDTAKVLMNNEAENAMHNTNTTIGSTILLQKNKVENHTLKYGTW 709
 62 AOMTPELOEITLVKLOLQALQONGSSVLSSEKSKRLTILNTMTSTYGTGVCNPDPQ 121
 DB AKTEPVSNFQNSTIKRIKIKVONVRAVLPNNEEYVQILLDMETTYSVANVCYTNG-- 767
 710 ECLLEPGLNEMANSIDYNERLWAMESRESEVQKOLPPEEYVVLAKNEKARANHYEDY 181
 DB ECLLEPGLNEMANSIDYNERLWAMESRESEVQKOLPPEEYVVLAKNEKARANHYEDY 181
 768 TCLSEPLTNIMATSRKYEELLMWMSKRDVGRALIPFPKYVDSNKLAKINGYSDA 827
 182 GDYMGDEYVNGVGYDYSRGQLIEDVETFEIKRPLYEHLAAVRAKIMAYPS-YISP 240
 DB GDSMRSSEYSDLE-----QDEKLYQELQPLTLNHAAYRSLRHRYSEYINL 877
 241 IGLPFAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIEAEKPFYVS 300
 DB DGPIDPAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIEAEKPFYVS 300
 878 DGPIDPAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIEAEKPFYVS 300
 DB DGPIDPAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIEAEKPFYVS 300
 301 GLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDEFLTAHHEMG 359
 DB GLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDEFLTAHHEMG 359
 938 GLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDEFLTAHHEMG 359
 DB GLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDEFLTAHHEMG 359
 360 HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPQEDNEHIN 419
 DB HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPQEDNEHIN 419
 998 HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPQEDNEHIN 419
 DB HIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPQEDNEHIN 419
 420 FLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGVVEPYPHET 479
 DB FLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGVVEPYPHET 479
 1057 FLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGVVEPYPHET 479
 DB FLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGVVEPYPHET 479
 480 CDPASLFVNSDYFIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIYQSKREAKLADA 1176
 DB CDPASLFVNSDYFIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIYQSKREAKLADA 1176
 1117 FDSKSTHVPANVYIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIYQSKREAKLADA 1176
 DB FDSKSTHVPANVYIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIYQSKREAKLADA 1176
 540 LRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV-STDMS 594
 DB LRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV-STDMS 594
 1177 MRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV-STDMS 594
 DB MRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV-STDMS 594

RESULT 8
 A34402
 pepitdyl-dipeptidase A (EC 3.4.15.1) precursor, testicular - rabbit
 N:Alternate names: angiotensin I-converting enzyme; dipeptidyl carboxypeptidase I; pepti
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 10-Sep-1999
 R:Accession: A34402; A60724; A56232; C18700
 R:Kumar, R.S.; Kusari, J.; Roy, S.N.; Softer, R.L.; Sen, G.C.
 J. Biol. Chem. 264, 16754-16758, 1989
 A:Title: Structure of testicular angiotensin-converting enzyme. A segmental mosaic isozy
 A:Reference number: A34402; MUID:9380303; PMID:2550457
 A:Accession: A34402
 A:Molecule type: mRNA
 A:Residues: 1-737 <KMD>
 A:Cross-references: GB:J05041; NID:g164744; PIDN:AAA31153.1; PID:g164745
 R:Sen, G.C.; Thekkumkara, T.J.; Kumar, R.S.
 J. Cardiovasc. Pharmacol. 16(Suppl. 4), S14-S18, 1990
 A:Title: Angiotensin-converting enzyme: structural relationship of the testicular and the
 A:Reference number: A60724; MUID:9155372; PMID:1705622
 A:Accession: A60724
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 73-173 <SEN>
 A:Note: Identical sequences were obtained for mRNAs from lung and testes
 R:Chen, Y.N.P.; Riordan, J.F.
 Biochemistry 29, 10493-10498, 1990
 A:Title: Identification of essential tyrosine and lysine residues in angiotensin convert

A:Reference number: A36232; MUID:91104959; PMID:2176870
 A:Accession: A36232
 A:Molecule type: protein
 A:Residues: 154-160;236-242 <CHE>
 R:Ikata, K.; Iai, C.Y.; El-Dorry, H.A.; Softer, R.L.
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
 A:Title: The NH₂- and COOH-terminal sequences of the angiotensin-converting enzyme iso
 A:Reference number: A90107; MUID:83048249; PMID:6291514
 A:Accession: C18700
 A:Molecule type: protein
 A:Residues: 33-35, 'SN', '38-39', 'SS', 'FAEL', '73' <TWA>
 A:Note: Several of the amino acids in reported are tentative
 C:Comment: The pulmonary and testicular isoforms of this enzyme differ substantially
 A:Superfamily: mammalian peptidyl-dipeptidase A
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; testis; transmembrane

Query Match 33.2%; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.8%; Pred. No. 1.6e-82;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

2 TIEBOAKTFLEKDNHEADLFYQSSLSAMNTNTTEENVOMNNAQKMAFLKEOSTL 111
 DB TDEAKANFVEEYEDTAKVLMNNEAENAMHNTNTTIGSTILLQKNKVENHTLKYGTW 134
 75 TDEAKANFVEEYEDTAKVLMNNEAENAMHNTNTTIGSTILLQKNKVENHTLKYGTW 134
 DB TDEAKANFVEEYEDTAKVLMNNEAENAMHNTNTTIGSTILLQKNKVENHTLKYGTW 134
 52 SAFLEKOSTIAOMYPILOEITLVKLOLQALQONGSSVLSSEKSKRLTILNTMTSTYGT 111
 DB SAFLEKOSTIAOMYPILOEITLVKLOLQALQONGSSVLSSEKSKRLTILNTMTSTYGT 111
 135 SAFLEKOSTIAOMYPILOEITLVKLOLQALQONGSSVLSSEKSKRLTILNTMTSTYGT 111
 DB SAFLEKOSTIAOMYPILOEITLVKLOLQALQONGSSVLSSEKSKRLTILNTMTSTYGT 111
 112 GKYCNPDNPOECLLEPGLNEMANSIDYNERLWAMESRESEVQKOLPPEEYVVLAKNE 171
 DB GKYCNPDNPOECLLEPGLNEMANSIDYNERLWAMESRESEVQKOLPPEEYVVLAKNE 171
 185 ANYCVRDYG--SCILPDPDLNLMATSRKYEELLMWMSKRDVGRALIPFPKYVEFTNK 242
 DB ANYCVRDYG--SCILPDPDLNLMATSRKYEELLMWMSKRDVGRALIPFPKYVEFTNK 242
 172 MARANHEYDGYDWRDYEYNGVGYDYSRGQLIEDVETFEIKRPLYEHLAAVRAKIMAY 231
 DB MARANHEYDGYDWRDYEYNGVGYDYSRGQLIEDVETFEIKRPLYEHLAAVRAKIMAY 231
 243 AALNNGYVADAGDSKRSKMTETPTLE-----QDEKLYQELQPLTLNHAAYRSLRH 292
 DB AALNNGYVADAGDSKRSKMTETPTLE-----QDEKLYQELQPLTLNHAAYRSLRH 292
 232 NAY-PSYISPTGLPFAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIE 290
 DB NAY-PSYISPTGLPFAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIE 290
 293 RHYGAQHINLEGPFAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIE 352
 DB RHYGAQHINLEGPFAHLGDMGRFWNTLSLTPFGQKNIDVTAMVDQAMDADRIE 352
 291 KEAEKPFYVSGLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDE 349
 DB KEAEKPFYVSGLPMTQGFENSMILTPGNOVAKVCHPTAMDLSG-DFTLMCTKYTMDE 349
 353 EEADEKFFISGLIPPEEYVVLAKNEKARANHYEDY 412
 DB EEADEKFFISGLIPPEEYVVLAKNEKARANHYEDY 412
 350 DEFLAHHEMGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPD 409
 DB DEFLAHHEMGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPD 409
 413 DLVYVHHEMGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPD 472
 DB DLVYVHHEMGHIQDYMAVAAPFLIRNGANGFHEAVGEIMSLAATPKHLKSLGILSPD 472
 410 FORDNETEINFLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGV 469
 DB FORDNETEINFLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGV 469
 473 -GGGYEHDINFLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGV 531
 DB -GGGYEHDINFLKQALITVGLPFTYMLKRWMMVFKGEIPKQDMKKNWEMKREITGV 531
 470 VEVVPEVDENYCDPASLFVNSDYFIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIY 529
 DB VEVVPEVDENYCDPASLFVNSDYFIRYRTYTRILYQFQFHEALCAAGHGTGPKCDIY 529
 532 CPPAPRSQDDEPDAFHFIPSSVPYIRYVSPILQFHEALCAAGHGTGPKCDIY 591
 DB CPPAPRSQDDEPDAFHFIPSSVPYIRYVSPILQFHEALCAAGHGTGPKCDIY 591
 530 TEAGOKLFLMLRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV 587
 DB TEAGOKLFLMLRLGSEPTLALENVGAKNMNVRPLNTEPEPLTWMKDNK--NSFYGV 587
 592 KEAGKRLADAMKLGYSKPPPEAMKYTGQPNMSASAMNYFRPLDMDLITENGREGKILG 651
 DB KEAGKRLADAMKLGYSKPPPEAMKYTGQPNMSASAMNYFRPLDMDLITENGREGKILG 651
 588 W-STDMSPYADS 599
 DB W-STDMSPYADS 599
 652 WQYTWTPNSARS 664
 DB WQYTWTPNSARS 664

RESULT 9
 S35484
 pepitdyl-dipeptidase A (EC 3.4.15.1) precursor, pulmonary splice form - rabbit
 N:Alternate names: angiotensin-converting enzyme; dipeptidyl carboxypeptidase I; kin
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 10-Sep-1999
 R:Accession: S35484; A23455; A18700; A38655; A49726; S17509

R:Thekkumkara, T.J.; Livingston II, W.; Kumar, R.S.; Sen, G.C.
 Nucleic Acids Res. 20, 683-687, 1992
 A>Title: Use of alternative polyadenylation sites for tissue-specific transcription of the
 A:Reference number: S35484; MUID:92178960; PMID:1311831
 A:Accession: S35484
 A:Molecule type: mRNA
 A:Residues: 1-1309 <THE>
 A:Cross-references: EMBL:X62551
 R:Itawa, K.; Blacher, R.; Soffer, R.L.; Lal, C.Y.
 Arch. Biochem. Biophys. 227, 188-201, 1983
 A:Reference number: A23455; MUID:84051289; PMID:6314908
 A:Accession: A23455
 A:Molecule type: protein
 A:Residues: 34-47; 'N', 49-55 <IMA>
 A:Experimental source: Lung
 R:Itawa, K.; Lal, C.Y.; El-Dorri, H.A.; Soffer, R.L.
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982
 A>Title: The NH₂- and COOH-terminal sequences of the angiotensin-converting enzyme isoform
 A:Reference number: A90107; MUID:83048249; PMID:6291514
 A:Accession: A18700
 A:Molecule type: protein
 A:Residues: 34-44; 754-755, 'L', 757 <IM2>
 R:Kumar, R.S.; Thekkumkara, T.J.; Sen, G.C.
 J. Biol. Chem. 266, 3854-3862, 1991
 A>Title: The mRNAs encoding the two angiotensin-converting isoforms are transcribed from
 A:Reference number: A38655; MUID:91139683; PMID:1847388
 A:Accession: A38655
 A:Molecule type: DNA
 A:Residues: 1-88 <KUM>
 A:Cross-references: GB:M58579
 R:Ramchandran, R.; Sen, G.C.; Misson, K.; Sen, I.
 J. Biol. Chem. 269, 2125-2130, 1994
 A>Title: Regulated cleavage-secretion of the membrane-bound angiotensin-converting enzyme
 A:Reference number: A49726; MUID:94124568; PMID:9294466
 A:Accession: A49726
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1236-1258 <RAM>
 A:Experimental source: testis
 R:Kiley, T.L.
 Biochem. J. 278, 375-380, 1991
 A>Title: The Mg(2+)-ATPase of rabbit skeletal-muscle transverse tubule is a highly glyco
 A:Reference number: S17509; MUID:91378880; PMID:1654880
 A:Accession: S17509
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 34-55 <KIR>
 C:Comment: This enzyme converts angiotensin I to angiotensin II in presence of divalent
 ver. the enzyme has been found also in renal tubules and intestinal mucosa.
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: alternative splicing; blood pressure control; chloride; glycoprotein; intest
 F:1-33/Domain: signal sequence #status predicted <SID>
 F:34-1309/Product: peptidyl-di-peptidase A, pulmonary #status experimental <MAT>
 F:59/79,150,222,448,512,680,698,717,945,1194/Binding site: carboxylate (Asn) (covalent)
 Query Match 33.2%; Score 1283; DB 1; Length 1309;
 Best Local Similarity 40.8%; Pred. No. 3.8e-82;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;
 2 TIEQAKTFLDKFNHAEDELFFQSSLASNNYNTITEE-----NVQNMNN--AGDKW 51
 Db 647 TDEAESRVEEEDRDFQAVNMEYANANNNTNTITEESKILLQKNQIAHHTLTYGWM 706
 707 -----ARRFDVSNFNQATSKRIKKVODLQRAVLPVKELEBYNQLILDMETISV 756
 52 SAEFKDOSTLAQMTPLQETQNTLVKLOALQOQSSVSEDKSKRLNTILMTSTIY 111
 Db 112 GKVCNDNDQECLELPEGLNEMANSNDNERLAMESRSVQKQLRPLVEEYVAKKE 171
 757 ANVCGRVGG--SCIQLEPDLTNLWATSRRKDELMLWTSWRDVGRIPLPEFPEYEF 814
 172 MARANHYEDYGDWGRDGYENGVDGYDYSRGQLIEDVETFEIKPLVEHLHAYRAKIM 231

Db 815 AARLNGYVADAGDSWRSMYETPTLE-----QDLERLEQLPLVYLNHAYVGRALH 864
 232 NAY-PSYISPIGICPAHLLGDMGRFNTNLSYTFEGQKRPNDVYDAMVDQAQRTF 290
 Db 865 RHVQAHINLEGPRTAHLGLNMAQWTSNIYDLVAPPPASSTMDATAMAKQWRRRF 924
 291 KEAEKFFVSVGLPMTQGEWNSMLDPGNQAVCAVCPHAWDLGK-DEIRILCTKYMD 349
 Db 925 EADKFFISGLLVPPEFNNKSKLEKPTGREGVYCHASAMDPYNGKDFIKCTVYME 984
 350 DELFAHHEMHIOYDAMVAAOPFLRNGAEGHEHVEGELMSLSATPHKLSIGLSD 409
 Db 985 DLVVHHEMHIOYFMOYKDLPAALRGANPGEHALGDLALSVSPHLHSINLSSS 1044
 410 FOEDNEFEINELKQALTIYGLPTTYMLKRWMEYFGEIPEDQMKWKEKREIVG 469
 Db 1045 -GGGEHDINPLKMALDKIAIFIPSYLVDEMRRVDSITENENQEWMSRLKYQGL 1103
 470 VEPYPHDEYICDPASLSEHVSNDSFIRYTRITLQFQFQALCOAKHSGPLHKCIS 529
 Db 1104 CPAPRSGDGFDPGAKFHPSVPIRYFVSFTIQFQFHALCKAAGHGPPLHTC 1163
 530 TEAGOKLFNMLRLGSEPTLALENVYGAKNMVRPLNYPEPLFTLKDN--KNSFVG 587
 Db 1164 KEAKRLADAMKLGISKPWEBAKVTTCGRNMSASAMNYFKPLMDMLLTENRG 1223
 588 W-STDWSPPYADQS 599
 Db 1224 WQYTWTPNSARS 1236
 RESULT 10
 S65472
 peptidyl-di-peptidase A (EC 3.4.15.1) precursor - horn fly
 N:Alternate names: angiotensin I-converting enzyme
 C:Species: Haemaphysa irritans (horn fly)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
 C:Accession: S65472; S65431
 R:Wijffels, G.; FitzGerald, C.; Gough, J.; Ridding, G.; Elvin, C.; Kemp, D.; Willadsen
 Eur. J. Biochem. 237, 414-423, 1996
 A>Title: Cloning and characterisation of angiotensin-converting enzyme from the dipe
 A:Reference number: S65431; MUID:96215437; PMID:8647080
 A:Accession: S65472
 A:Molecule type: DNA
 A:Residues: 1-611 <WJD>
 A:Cross-references: EMBL:L43965; NID:9908759; PID:AAA70427.1; PID:9908760
 A>Note: the source is designated as Haemaphysa irritans exigu
 A:Accession: S65431
 A:Molecule type: protein
 A:Residues: 18, 'P', 20-42; 75-77, 'Q', 79-81, 'X', 83-84; 179-190 <IMW>
 A>Note: the source is designated as Haemaphysa irritans exigu
 C:Gene: ACE
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: glycoprotein; metal binding; peptidyl-di-peptidase hydrolase; zinc
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-611/Product: peptidyl-di-peptidase A #status predicted <MAT>
 F:53,196,331/Binding site: carboxylate (Asn) (covalent) #status predicted
 Query Match 27.3%; Score 1054.5; DB 2; Length 611;
 Best Local Similarity 37.7%; Pred. No. 1.6e-66;
 Matches 222; Conservative 105; Mismatches 229; Indels 23; Gaps 8;
 1 STIEQAKT-FLDKFNHAEDELFFQSSLASNNYNTITEE-----NVQNMNN--AGDKW 59
 Db 18 ATKEIVATETVLQINNELAKHTVWEVSAVSNNTDNERLRNLSISAKFKLEVA 77
 60 TLQAMYPLQETQNTLVKLOALQOQSSVSEDKSKRLNTILMTSTIYVCPDN 119
 Db 78 KDIOGFNNKRTGSDVDRQFSLKSTGYSALPAEDVLELVISAMESNRAKRVCDYK 137
 120 PQEC-LLEPGLNEMANSNDNERLAMESRSVQKQLRPLVEEYVAKKEANHY 178

Db 138 SACDLSLDEIEIITKSDPEELKYWTQYDKAGTPTRSNFEKVELNITKSAKLNNF 197
 QY 179 EDYGDYWRGDEYVNGVDYDSRGOLIED-VENHTFEIKPLYEHLHAYVRAKLNNAY-DS 236
 Db 198 TDCAEWLDEYE-----DAFFEDQLAIFEDIKPLQVGVYARLNRKRYGDE 246
 QY 237 YISPIGCLPRLHLLGDMKGRWTLKSLTVFGOKPNIDVTDAVQADAMQRIEKAERF 296
 Db 247 VVSKTGPLPMLHLLGNMAAQWSSIAIVSPFPEKPLVDVSDKVAQGYPLKMFQGDGF 306
 QY 297 FVVGKLPNMTQGFENSMILDPGNVQKAVCHPTAMLG-KGDFRILMKRYMDDPLTAH 355
 Db 307 FQSMGLAKLPEQEWDSILEKPDGDLCHASAMFPYTLTDVRIKQCTRYVQDQFPTVH 366
 QY 356 HENGHIQYDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNE 415
 Db 367 HEMGHIOYFLOYOHOFVYVYTGANGPFHEAVGDVLSVTPKHLEHVGLK-NYSDNE 425
 QY 416 TEINPELLKQALITVGLPFTYMLEKRRMVFGEITKDDQMKMKWEMKEIYGVVEPVH 475
 Db 426 ARINQLFLALDKIVLPFAFTMDKRYMALFRQADKSMNCALFWLREESGIEPPVVR 485
 QY 476 DETYCDPASLFEVSNDSYFIRYTRTYLQFOFQALCOAA-----KHGGLPKODISNS 529
 Db 486 TEKDFAPAKYHVSADVEYLRVYSTITQFQFKSACTAGERYVQTEYPLDNCDDYGS 545
 QY 530 TEAGOKLFNMLRIGKSEPTLALENVYGAKNMNVRLNYFEPPLTWLK 578
 Db 546 KEAGKLFENMLSLGASKPMPDALLFAVNGERTMGKALAHYFELRWLE 594

RESULT 11

angiotensin-converting enzyme-related protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 21-Jul-2000
 C:Accession: J05374
 R:Taylor, C.A.M.; Coates, D.; Shlirsa, A.D.
 Gene 181, 191-197, 1996
 A:Title: The Ace2 gene of Drosophila codes for an angiotensin-converting enzyme homolog
 A:Reference number: J05374; MUID:97128790; PMID:8973330
 A:Molecule type: mRNA
 A:Residues: 1-630 <TRAY>
 A:Cross-references: EMBL:X96913; NID:g1405881; PIDN:CAA65632.1; PID:g1405882
 C:Genetics:
 A:Gene: Ace2
 C:Superfamily: mammalian peptidyl-di-peptidase A

Query Match 26.6%; Score 1027.5; DB 2; Length 630;
 Best Local Similarity 36.0%; Pred. No. 1.4e-64;
 Matches 215; Conservative 110; Mismatches 250; Indels 23; Gaps 9;
 QY 6 QAKTFLDKRNEHAEDLFTQSSLASWNYNTNTEENVQNMNAGDWSAFLKQSTIAQ 65
 Db 33 EARREFELENEQRLRRFHEFLSGVNYNTNTEANRQAMIEVYANNAELNRLOAQTSS 92
 QY 66 PLQELQNLTVKIQLOALQONGSSVYSEDKSKRLNTILMTGSTYTGVCVCPNDQEC-L 124
 Db 93 DYVQSEDADIRQOAHSLKLSALNADYTLALQNAISSMOTNATATVCSTYNSDCSL 152
 QY 125 LLEPGLNETIMANSLDYNERLWAMESRSEVGKQLRPLVEEYVVLKNEARANHEDYD 184
 Db 153 TLEPHIGRLSHSRPALAWYRMWHDKSGTPMKNQFAEYVRLTRKASQULNGHRSYDY 212
 QY 165 WRGDEVNGVDGYDSRQOLLEDEVEHTFEIKPLYEHLHAYVRAKLNNAY-PSYISPIGC 243
 Db 213 WVOFEY-----DPPFER-----QLDATEKQLPLTRYRQHGYYRFLRQYBGDVAPAGN 262
 QY 244 LPAHLIGDMGWFNMLSLVVPFGOKPNIDVTDAVQADAMQRIEKAERFVYVGLP 303
 Db 263 IPISLIGNMGOSNELLDTFTPEKPFVYVKAEMERQGYVOKLFEIGDQFQFOSLGM 322

QY 304 NMTQGFENSMILDPGNVQKAVCHPTAMLG-KGDFRILMKRYMDDPLTAH 362
 Db 323 ALPPEFWNLISVTLRPDD-RQVCHASAWDFQDSVRIKMCLEVDSHFYVHHLGHQ 381
 QY 363 YDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNETEINPL 422
 Db 382 YLYQEQGPVAVYRGAPNFGFHEAVGDVLTALYSKHLKALGLE-NORDEKSKINLE 440
 QY 423 KQALITVGLPFTYMLEKRRMVFGEITKDDQMKMKWEMKEIYGVVEPVHDEYCDP 482
 Db 441 KQALSKIVYLPFGYAVDXYRAVAFERNELDESQWNGFQWMSRSEGVPEPVFRTKEDRP 500
 QY 483 ASLFVSNDSYFIRYTRTYLQFOFQALCOAAKEGP-----LHKDINSREDAQKL 536
 Db 501 PAKYIHDADVELIRFPAHIFQFQHKVLCRAAGYABNRSRLTLDNCDDIGSKAAGSL 560
 QY 537 FNMRLIGSEPTLALENVYGAKNMNVRLNYFEPPLTWLKDQNKNSFVSTWDSF 594
 Db 561 SQFLSKGNSRHKEVLEFTEGTENDPALLEYFEPYOWLKOE-NSRLGVPLGMGP 616

RESULT 12

peptidyl-di-peptidase A (EC 3.4.15.1) 67k precursor - fruit fly (Drosophila melanogaster)
 N:Alternate names: angiotensin-converting enzyme
 C:Species: Drosophila melanogaster
 C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 26-Feb-1998
 C:Accession: A57533
 R:Corneil, M.J.; Williams, T.A.; Lamango, N.S.; Coates, D.; Corvol, P.; Soubrier, F.;
 J. Biol. Chem. 270, 13613-13619, 1995
 A:Title: Cloning and expression of an evolutionary conserved single-domain angiotensin
 A:Reference number: A57533; MUID:95293950; PMID:775412
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-615 <COR>
 A:Cross-references: GB:U25344
 A:Gene: Flybase:Ance
 A:Cross-references: Flybase:FBgn0012037
 C:Superfamily: mammalian peptidyl-di-peptidase A
 C:Keywords: peptidyl-di-peptide hydrolase

Query Match 26.4%; Score 1022; DB 2; Length 615;
 Best Local Similarity 35.6%; Pred. No. 3.3e-64;
 Matches 213; Conservative 118; Mismatches 245; Indels 22; Gaps 9;
 QY 4 EEOAKTFLDKRNEHAEDLFTQSSLASWNYNTNTEENVQNMNAGDWSAFLKQSTIAQ 63
 Db 22 EIOAKTFLDKRNEHAEDLFTQSSLASWNYNTNTEENVQNMNAGDWSAFLKQSTIAQ 81
 QY 64 MYPLQELQNLTVKIQLOALQONGSSVYSEDKSKRLNTILMTGSTYTGVCVCPNDQEC 123
 Db 82 KFORRSYQSEELKQFALTKLGYALPDDYAEALDTLSLAMESNEFKVAVCYKQSTK 141
 QY 124 LLEPGLNETIMANSLDYNERLWAMESRSEVGKQLRPLVEEYVVLKNEARANHEDYD 182
 Db 142 DLALDPELEIYISRSRDEELAYWRFYQKAGVAVSQERVYELNTRAKAKLNFTSCA 201
 QY 183 DYWGQDEVNGVDGYDSRQOLLEDEVEHTFEIKPLYEHLHAYVRAKLNNAY-PSYISPT 241
 Db 202 EAWLDEYE-----DFFEQQLEDI--PADIRPLLPADPWLCAFLRKHNYDAVAVSET 251
 QY 242 GCLPFAHLIGDMGWFNMLSLVVPFGOKPNIDVTDAVQADAMQRIEKAERFVYVGLP 301
 Db 252 GPIFMHLLGNMNAQWSEIDIVSPFPEKPLVDVSAEMEQATPLKMFQMGDDFTTSKN 311
 QY 302 LPAHTQGFENSMILDPGNVQKAVCHPTAMLG-KGDFRILMKRYMDDPLTAH 360
 Db 312 LTKLPRDQFPMHLLGNMAAQWSSIAIVSPFPEKPLVDVSAEMEQATPLKMFQMGDDFTTSKN 371
 QY 361 IOYDMAVAAPPELLRNGANGFHEAVGEIMSLAATPKHLKISGLSPDQEDNETEIN 420

Db 372 IOEFLQYQHOPFYRTGANGPHEAVGDLVSLSTVSPKHLKIGILK-DYVRDEBARINQ 430
QY 421 LKQALTYGTLPFTYMLEKRWMTFKGIPDPQMKMKWEKREIVGVVEVPDEHYC 480
Db 431 LFTLALDKLTVLPFAETMDKRYMSLFRGEVDANNKCAWKLDETSGETPVVSEDF 490
QY 481 DPASLPHVNDSEFIYRTFLYQFOFQALC-QAKH-----EGPLKCDISNSTEQ 534
Db 491 DAPAKHISADVEYELKLYVSLIQOFYKACIKAGQYDPDVELPLDMDCLYSGARGA 550
QY 535 KLENNLRICKSEPTWLTALENVGAKNNVPLNTPEPLTYLKQON-KNSFVGWST 590
Db 551 AFNNMLSMGSKRPWDALFAFNGERIMSGKATAEYFEPLRWLLEAKENKNNVHIGMTT 608

RESULT 13

T15792

hypothetical protein C42D8.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T15792

R:Hallsworth, K.

Submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid C42D8.

A:Reference number: Z18405

A:Accession: T15792

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1,907 <NAL>

A:Cross-references: EMBL:056966; NID:g1293844; PID:g1293847; PIDN:AAA98719.1; GSPDB:GN00

A:Experimental source: strain Bristol N2; clone C42D8

C:Genetics:

A:Gene: CESP:C42D8.5

A:Map position: X

A:Insertions: 140/3; 194/3; 300/2; 467/3; 551/2; 600/2; 697/3; 774/2; 851/3

Query Match 16.5%; Score 638.5; DB 2; Length 907;

Best Local Similarity 27.0%; Pred. No. 7.5e-37;

Matches 167; Conservative 132; Mismatches 280; Indels 39; Gaps 18;

QY 4 EBOATPTLDKFNHEADLFYQSSLASMYNTNTEENVOMNNAAGKWSAFLKEOSTLQ 63
Db 178 EEKIRSWLAGYEADAIKLVREVALSGWRFPNDASPSIKALDAENVLTVEVSTSMQK 237
QY 64 MYPLQEIQNTLVKLOLQALQONGSSVLSDESKRLTILNTMSTIYSTGKVCNPDNPOEC 123
Db 238 QPDMASVTDEKVRKOLGYVEEGMSALAPSRFADYSQAQALNRDSDSTICDKDVPPEC 297
QY 124 LLEPGINETMANSLDYNERNLMAWESRSEVQKOLRLYEYVLKNEAARAHYHDYGD 183
Db 298 ALQKIDSDSIFRREKASRLQHLWVSIVTAIAKS-KFSYNNITITISNEGAKLNGFANGGA 356
QY 184 YMRGDYEVNG-VQGYDYSRGQLEDEYHEETKEIKPLYEHLHAYVRAKIMNAV--PSYIS 240
Db 357 MMRSAFPMSSKVKHAEF---DLNKQIDKISTYTOPFQOLHAHMRQOLAGIYSNPGLSK 413
QY 241 IGCPLPAHLGDMGREFNTNLYSLTFPGQKPNIDVTDAVD---QAMDQRIFKEAEK 296
Db 414 DGPFAHLFGSLDGGMSAHYEQTKEPPEES--ETPEAMLSAFNTQYITTKKMFVAYRY 471
QY 297 FVSVGLPNNMNOGFENESMLDPPGVQKAVCHP-TAMDL-GKDPRLIMCKYVMDPLTA 354
Db 472 FKSAGFPHLPKSYTSSIFARWVS-KMICHPAALADWRAPNDFRYAACQOLDEPPEQA 530
QY 355 HHEGHIQYDMAVAAQFPLLRNGANEGFHEAVGEIMSLSATPKHLKISGLSPDEQDN 414
Db 531 HSLVQYTYQVLYDQSLFRQASPVITDAIANAFHLSSTNPHYLLSQGLVSEHLDIK 590
QY 415 ETE-INFLKQALTYGTLPFTYMLEKRWMTFKGIPDPQMKMKWEKREIVGVVEPY 473
Db 591 DSVIINKLYKESLSEFTKLPFTIADMMWRELEFGIVPKKLNDRMWEINKKEGVASPO 650
QY 474 PHDETQCDPASLFRH-VSNDSYFIRYTRTL---YQFOFQALCQAAA---KHGGLKHC 524

Db 651 PYNTSND--ALHNSVSQVHS---PATRLLISVYLFOILKALCORELFWLSEG-----C 701
QY 525 DISNSTAGOKLFNMLKSEPTWLTALENVGAKNNVPLNTPEPLTYLKQON--582
Db 702 ILSEDTT--EKLRRTMLGSSITWKLALMLISGKGELDAQPLLEYEPLINLNRINEID 759
QY 583 NSFVGSWTDSMPYADOST 600
Db 760 QVVVGMDGEPTFVEEI 777

RESULT 14

C83696

hypothetical protein BH0371 [Imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: C83696

R:Yakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: C83696

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-532 <STO>

A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BAB04090.1; GSPDB:G

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0371

Query Match 4.1%; Score 157; DB 2; Length 532;

Best Local Similarity 21.1%; Pred. No. 0.0034;

Matches 118; Conservative 83; Mismatches 213; Indels 144; Gaps 29;

QY 4 EBOATPTLDKFNHEADLFYQSSLASMYNTNTEENVOMNNAAGKWSAFLKEOSTLQ 59
Db 3 EODIRFLSEQKRVEDLYQVYLNHNVAVTTGEQMSDKHQSLSYMAHFSRSEFOK 62
QY 60 -----TLAOMYPLQEIQNTLVKLOLQALQONGSSVLSDESKRLTILNTMSTIY 109
Db 63 VTRFKRISLPLMQRRQDLHDHMKIKNOFE--EGTRQILSTLE--KKISHVFTFPQOV 118
QY 110 SNGKVCNPDNPOECLEPGLNETMANSLDYNERNLMAWESRSEVQKOLRLYEYVLK 169
Db 119 NSGRVSNNE-----LIDLIRYDLDEHRRQAWFA-SKEVGKTEKOLLQILKKR 166
QY 170 NEMARANYEDGYDGRDYEYNGVDYDYSRGQLEDEYHEETKEIKPLYEHLHAYV 226
Db 167 NEVARNIGEPF-----YHMSFAQGEIDLEQTFAMFETIKSSDQAFPMI 211
QY 227 -----RAKLMNAVSTISPIGCLPAHLGDMGREFNTNLYSLTFPGQK-PNIDVTDA 278
Db 212 KDEIDEEKRAKVLKIKKDLRP-----WDVYDPFQDAPASIEHVD- 250
QY 279 MYDQAMDQRIFKEAEKFEVSVGLPNNMNOGFENESMLDPPGVQK-AVCHPTAMDLGKD 337
Db 251 -FDSYTKDQDLQVYVSQFOAMELP--IDDLIKSDILPRKKNKPFQFC--TDMD-RRGD 304
QY 338 FRIIMCTKVYMDPFLAHHEMGH-IQYDMAVAAQFPLLRNGANEGFHEAVGEIMSLSAT 396
Db 305 IYVLLNDQSMYVAVALLHEFGHAYYFKFIDRLPFLR-----FH-----SHLTLT 351
QY 397 PHLKISGLLS--PDQF-----DNET-----ELNFLKQALTYGTLPFTYMLEKRW 443
Db 352 EASALFFGRTKMAWYERFLGIDRETERIGRNNKMLQROM-VVST-----RW 400
QY 444 NY-----EIPKQDMKKMKWEKREIVGVVEPVPDEYCDPASLFRH-----N 490
Db 401 MLTFSEFSKSLYEDPDQDINALMKLVKEIQTMAP--PEDIGSPMAKMHFSLAPVYQ 458
QY 491 DY-----SFRYYTRT 501

Db 459 DYLLGMAASQLHHYIKT 476

RESULT 15

AF1310

probable thermostable carboxypeptidases homolog lmo186 [imported] - *Listeria monocytogenes*C/Species: *Listeria monocytogenes*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C/Accession: AF1310

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AF1310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99964.1; PID:916411339; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo186

C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match

Best Local Similarity 4.0%; Score 154; DB 2; Length 502;

Matches 128; Conservative 101; Mismatches 215; Indels 192; Gaps 35;

2 TTEOAKTFIDKFNHEADLFYQSSLASWNT-----NITEENYONNNA 47

4 TLEEFYAYIKKMALEAL-----ALVYMDLRTGAPAKMGESRDVIGVSELEFNQTS 59

48 GKKSAFLKEOSTLAOMPLOEIONLYKLOALQONGSSVLSDESKRL-----N 99

60 -EEMAFIAGLN-----QDENIS-EITRTLEE--SOKTYDKNKKIPKEVAYEYK 106

100 TLNMTSTYSGKVCNPNPOECILLEGNEIMANSIDYERLWAMESRSEVKQRL 159

107 KLVAAQETAMTARSON-----DFAFEPLTKIL----- 136

160 PLYEEVVLKEMARANHEDYGDYRGDYEVNGVGY--DYSRQLIEDVHEFEIKP 217

137 -----EMKR-----KFEYEW--GYEENKIDTLIDYEPGVTVSLDSVEKVR- 177

218 LYEHLAAVYRAKIMNAYPSYISPIGCLPAHLIGDMGGRWTNLSLYVFGOKPNIIDVTM 277

178 --DGIMA--IREKIENE-----EAGRLDETV-HPEATGLNTGDR 200

278 AMVDAMDAQRIFKEAEKFEVSVGLPNTQGFWENSMLTDPGNVOKAVCHPTAMD LGKD 337

201 TKISRA-----KOKESIRI-LNKKGFDF-----EAGRLDETV-HPEATGLNTGDR 243

338 FRLMCTVYMDDELTA-----HHEMGHIQY-----DMAVAAOPFLLRNGANGFHEA 387

244 VRI--TTRYNENDFKAAVFGTHIEGHAIEQNPAALVGP--LANGASMGIHESQSLF 299

368 -EIM--SLSAATPRHLKSGILSPDQ-----ED-----NETEINFLKQALTYVGLP- 433

300 YEIIIGSSIAFWKSNVADFOAITKPAFDQVKLEDFYRAVNISESSLRIEADTL--TV 357

434 ---FTYMLEKRWAVFKEIPRDOMKKWMEKREIVGVPPVPHDEYCPASLFFVSN 490

358 HIMIRYELER--ALINGELEVKDLPKAMGKYBEYIGI--RDNDINGVLODIHNAAG 411

491 DYSFIRYTRTL-YOFQOALCOAAKHEGPHLKCDISNSTEAGOKLF--NMLRLGK 546

412 DGYEYPSYALGLMTAAQFFNQM--QKEIPNIDAILASDYSSELTWLTGKHTFGK 467

547 PWTTLALENVGAKKNNVAPLINTFEPPLTWLKDOK 582

Db 468 KPEILIDPTG-EGINPTYLIDLEKRYAYVQENK 502

RESULT 16

AE1682

probable thermostable carboxypeptidases homolog lin199 [imported] - *Listeria innocua*C/Species: *Listeria innocua*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C/Accession: AE1682

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AE1682

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-502 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97229.1; PID:916414500; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin199

C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match

Best Local Similarity 3.8%; Score 147; DB 2; Length 502;

Matches 124; Conservative 95; Mismatches 223; Indels 196; Gaps 33;

2 TTEOAKTFIDKFNHEADLFYQSSLASWNT-----NITEENYONNNA 47

4 TLEEFYAYIKKMALEAL-----ALVYMDLRTGAPAKMGESRDVIGVSELEFNQTS 59

48 GKKSAFLKEOSTLAOMPLOEIONLYKLOALQONGSSVLSDESKRL-----N 101

60 -EEMAFIAGLN-----QDENIS-EITRTLEE--SOKTYDKNKKIPKEVAYEYK 109

102 LNTMTSTYSGKVCNPNPOECILLEGNEIMANSIDYERLWAMESRSEVKQRL 161

110 AQA-ETAMTARSON-----DFAFEPLTKIL----- 136

162 YEEVVLKEMARANHEDYGDYRGDYEVNGVGY--DYSRQLIEDVHEFEIKP 219

137 -----EMKR-----KFEYEW--GYEENKIDTLIDYEPGVTVSLDSVEKVR- 177

220 EHLAAVYRAKIMNAYPSYISPIGCLPAHLIGDMGGRWTNLSLYVFGOKPNIIDVTM 279

178 --DGIMA--IREKIENE-----EAGRLDETV-HPEATGLNTGDR 202

280 VDAQMDAQRIFKEAEKFEVSVGLPNTQGFWENSMLTDPGNVOKAVCHPTAMD LGDPR 339

203 ISEA-----KOKESIRI-LNKKGFDF-----EAGRLDETV-HPEATGLNTGDR 245

340 LLMCTVYMDDELTA-----HHEMGHIQY-----DMAVAAOPFLLRNGANGFHEA 385

246 I--TTRYNENDFKAAVFGTHIEGHAIEQNPAALVGP--LANGASMGIHESQSLF 301

386 -----VGEINISAAATPRHLKSGILSPDQ-----ED-----NETEINFLKQALTYVGLP- 432

302 IIGSSIAFWKSNVADFOAITKPAFDVH-----LEDYRAVNISESSLRIEADTL--TV 355

433 P---FTYMLEKRWAVFKEIPRDOMKKWMEKREIVGVPPVPHDEYCPASLFFVSN 488

356 PLHIMIRYELER--ALINGELEVKDLPKAMGKYBEYIGI--RDNDINGVLODIHNAAG 409

489 SNOYSFIRYTRTL-YOFQOALCOAAKHEGPHLKCDISNSTEAGOKLF--NMLRLGK 544

410 GGDGYEYPSYALGLMTAAQFFNQM--QKEIPNIDAILASDYSSELTWLTGKHTFGK 465

545 SEPWTTLALENVGAKKNNVAPLINTFEPPLTWLKDOK 582

Db 466 TKKPLELIDTNG-EGLNPTYLILDLLEKRYAVYQENK 502

RESULT 17

peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)

C:Accession: A12011

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: A12011

R:Kanezo, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, R.; Kanezo, T.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A:Reference number: AB1807; MID:21595285; PMID:11759840

A:Accession: A12011

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-987 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA078013.1; PID:g17135467; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all1647

Query Match 3.6%; Score 139.5; DB 2; Length 987;

Best Local Similarity 19.6%; Pred. No. 0.15;

Matches 117; Conservative 80; Mismatches 208; Indels 191; Gaps 26;

QY 34 TTTTENVQNNANAGDKMSAFLEQSTLAQMYPLQIQTQNYVQLQALQONSSVLSLD 93
 Db 106 TLTTEHQVSOOLVACVQEAFLPLOTVMFLDEGEPLGIQSLT-----QITAEWQS-LSDD 159
 QY 94 KSKRLNTLTMTSTIYTGKVCNPDN-----POECLLEPLGLNEMANSID 139
 Db 160 ALELCNSPDLMTYITTSSTGRKGMNLNHRGMRLTMQNTFSLQGDRAQRTSIC 219
 QY 140 YNERLWAMESRSEVQKOLRPLVEEYVVLKEMARANHEDYGDYRGDEYVNGDGYIY 199
 Db 220 FDISVME-IPWTLMSGATTCPCVOREVVLNMFEPAR-----WIOETQIVMHFVVS 268
 QY 200 SRQGLIEDVHEH-----TPEEIKPLVEHHAAYRAKIMNAPSYISPICSLPAHLIGDMGR 256
 Db 269 LRFGEFISALENETWSFQWLMWPSGA-----LPHSFT-----QR 304
 QY 257 W-----TNYSLTVPFGQKPNIDVTDAVQAMDQ-----RIFEAKKFFVS- 300
 Db 305 WIDRHGLKTGLANLYGPT-----EASIDVYCHLITERPDERLTQIPIDGKADINYYVKYL 359
 QY 301 --GLPNTQG-----FWNSM-----LTDEGNVOKAVCHPTAMDICKGD--ERILMCTVY 346
 Db 360 DGGMQPQPGNMGELWLGVOALAGYLKDEKTAQAFCPNPFTDI--PGDIYTTGDLVKE 418
 QY 347 TMDFTLAHHEMGIHQDMAVYAAOPFLIRANGANEGEHVAEIMSLSATPKILKISGLL 406
 Db 419 LPQDTIYH---GALIDQVRI-----RGFIELGEIESVLTTHP--DYREAAAL 462
 QY 407 SPDFQEDNETEINFLKQALTYGTLPFTYMLEKRWMTFKGELPKDQMKKWKWEKREI 466
 Db 463 AVDGEQKRLV-----ACISGKIKIKRKFKEYLEOK--- 494
 QY 467 VGVVEPRPHDETYCDPASLFHVSNDSYFIRYTRTLVQFOFQALQAANHEBRLHKCDI 526
 Db 495 -----LPH---YIIP-----ORFLMLDSL--PKHNGKLDKRAL 523
 QY 527 -----SNSTGAGKLFNMLRIGKSEPTLALLENVVGAKNNVBPILLNYEPLETLW 577
 Db 524 VTQLTSDSPSSPSPLPLPLGPRQNR-----LVVYEPPEPYQWL 562

RESULT 18
 B82938
 zinc metalloproteinase oligoendopeptidase F UU065 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: B82938

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Caswell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a

A:Reference number: A82870

A:Accession: B82938

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-608 <GLA>

A:Cross-references: GB:AB002106; GB:AF222894; MID:96899011; PIDN:AAF30470.1; GSPDB:GN

A:Experimental source: serovar 3; D1ovar 1

C:Genetics:

A:Gene: pepF-1; UU065

A:Genetic code: SGC3

Query Match 3.6%; Score 139; DB 2; Length 608;

Best Local Similarity 19.1%; Pred. No. 0.079;

Matches 132; Conservative 93; Mismatches 226; Indels 240; Gaps 32;

QY 4 EEOAK-----TFIDKFNHEAE-----DLFYQSSIASMYNTNTEENVQNMN 45
 Db 28 EKKAKYIAFPFPLDSKQNFQMOQILBEFTIVANRFY--NYVSNLNTNVVD----- 78
 QY 46 NAGDKWSAFLEQSTLAQMYPLQ-----EIQNLTVKIQALQALQONSSV- 89
 Db 79 --PKMNSW--SOKLSASFYELETAISNYSNVLANEAKIKELYLDSOLNVTYTKRYEIF 133
 QY 90 -----LSEDKSKRLNTLTNT--MTSTYST-----GVCNPDNPOCLL 125
 Db 134 RYQPTLNNBOSKLESTILRADEGSTIFSTYTNNDKFSADAKGKPKHKEALAFV 193
 QY 126 LEPGLNEMANSIDYNERLWAMESRSEVQKOLRPLVEEYVVLKEMARANHEDYGDY 185
 Db 194 HLKSDRYLRKSA-YLSMYRAYYDSRESITK--LYYNYLSL--NQAKKANFDY--IA 246
 QY 186 RQDYEYNGVDYDYSRGQIEDVEHTPEEIKPLVEHHAAYRAKIMNAPSYISPICLP 245
 Db 247 KRAFP-----DVAQKSLITLIDQVLYKDTNEDKRYKANTYLLKILKY--SKIP----- 295
 QY 246 AHLIGDMGRFETNLYSLTVPFGQKPNIDVTDAVQAMDQAFKEAKKFFVSGLPNM 305
 Db 296 -----WDNGLPL--ISKKIDIPREAK-QMADLSLIGDE--YVS-----NI 333
 QY 306 TGGFENSMILDPGAVQAVCHPTAMDICKGDFRILMCTKYTMDFTLAHHEMGIHQDM 365
 Db 334 KRAFPDEKWSMLPQKRGAGAYSIGTGKISKYYILMNTNSLRDYOYIVHELGHSMHS- 392
 QY 366 AVAADPFLIRANGANGFHE--AVGEIMSLAATPKHKLKISGLSPDFQEDNETEINFL- 421
 Db 393 -----LYSNRTQKITSDYKIFFAELIASIS-----NEYLYNYL 425
 QY 422 -----LKQALTYGTLPFTYMLEKRWMTFK-----GEIPDQMKKWE 461
 Db 426 LEKYKNDLEMKIMLIDEMISGFPAVTRQVIFSNFEMIANELINGAFADVAVKEQYK 485
 QY 462 MKREIVG--VYEP-----YPHDTYCDPASLFHVSNDSYFIRYTRTLVQF 505
 Db 486 LELETTYNRPVIEDLNSIYSLSSITPLRLPH-----TYGVNFY-----YKY 526
 QY 506 OFQFALCQAKKEGFLHKCDISNSTGAGKLFNMLRLG-----KSEPW 548
 Db 527 AVGVAAIAISGR-----VFTKYVGAKQKAYDFELSSGSGNDPDTITIKLGVLDLTKPQAW 580
 QY 549 TLALENVVGAKNNVBPILLNYEPLETLW 579
 Db 581 OEALTEIV-----LWIKD 593

RESULT 19
 D82881
 zinc metalloproteinase oligoendopeptidase F UU521 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C/Accession: D82881
R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A/Description: The complete sequence of *Ureaplasma urealyticum*. Alternate views of a mit
A/Reference number: A82870
A/Accession: D82881
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-611 <GLA>
A/Cross-references: GB:AE002150; GB:AF222894; NID:g6899515; PIDN:AAF30934.1; GSPDB:GN001
A/Experimental source: serovar 3; biovar 1
A/Genetics:
A:Gene: pepF-2; U0521
A:Genetic code: SGC3

Query Match 3.5%; Score 136; DB 2; Length 611;
Best Local Similarity 19.5%; Pred. No. 0.13;
Matches 119; Conservative 100; Mismatches 256; Indels 134; Gaps 27;

QY 31 NYNTNTEENVONMNNAGDKWSAFK-ESSTLAOMYPLQEIOLNLYK----- 76
DB 69 NYVSNKLQTNL-IDNEMLAMSOKIEHQHRAKIF--INFENIAKKNLDINSYIKNS 124
QY 77 -----LQDALQOQSSVYSEDKSKRLNTILNTMTST-----YSTG 112
DB 125 LKRYOLEYELMWEKHEKHIIEQOQKVYTAISRESSFGDIFVLLDSMDQDGINYKQ 184
QY 113 KVCNPNPOBCLLEPGLNEIMANSLDYNERLWAMSRSSEVSKOLRPLVEEYVUKNEM 172
DB 185 KVCFKK--OTDLVASKSDRALRKSAYESHRAIYDILNRTSKL---LYEE-VKQNEL 238
QY 173 ARAHHEDYGDVWRKGDYEVNGVDYDSGOLLEDEHTFEELKPLYEHLAAVVR--AKL 230
DB 239 AKLHNFKDY-----ISADAFSDKVDKMF-----INHIYQTKFAFG 275
QY 231 MAAVPIYISPIGLPAHLGLDMGKRWTLISLTFVFGKPNIDVTDA--VQAMDQRI 289
DB 276 INRYTYRT-----LELKOQYOLTKVEPWDMKMLDIDKKNMESIESAKML 320
QY 290 FKAKEFVSVGLPNTQGFENSMMLTDCGNOKAVCHPTAMDLDGDRILMCTKVTMD 349
DB 321 TLEALLLGSSEYINVOAFNEQWIMMNNKISAYISISNTKGLDKFIILMNDETYN 380
QY 350 DFLTAHHEGHIOYDMAYAAQPLPLRNGANGEGHEAVGEIMLSAATPHKLSIGLSPD 409
DB 381 SILTLVHELGHSHV--TYFA-----NOSQEVYNEVETFYAEIASITNELIMNYHLK-K 431
QY 410 FQDNTEIEINFLKQALTYGILPFTYMLEKRWMTYKGEIPKQDMKKWKKREYV 468
DB 432 YENDDLMLRLYILDEMIISGFLATTTTQALFISNEMVA-----NEWINQEEFSNMKIVL 484
QY 469 VVEPVPHDET-----YCDPASLPHVSNDYSFIRYRTTLQFOFO--EALQOAK 516
DB 485 AYLEINHDTGYKYNKKNKISKYDEANALLINIPH---FYISNFYVYKVGQIC----- 536
QY 517 HEGPLKCDI--SNSTAGOKLFNMLRIGS--EFTWLTALENVGAKNNVRLPLNTEPL 574
DB 537 --GLINAIRIFNNKANAKEYCFCKSGSLSP--LETINILIDIK--INENDWEEVNIIF 591
QY 575 -TWLKDQNK 582
DB 592 NSMIDDIYK 600

RESULT 20
D69943
carboxypeptidase homolog ypwA - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C/Accession: D69943
R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
lechi, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Kocher, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogilwara, A.; Oudga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portele
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani
A/Authors: Schleicher, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: D69943
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-501 <KUN>
A/Cross-references: GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB14125.1; PID:g26346
A/Experimental source: strain 168
C/Genetics:
A:Gene: ypwA
C:Superfamily: Thermus aquaticus carboxypeptidase Taq

Query Match 3.5%; Score 135; DB 2; Length 501;
Best Local Similarity 20.9%; Pred. No. 0.11;
Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

QY 162 YEEYVVL--KNDMA--RANHVEDG-----DVRGDYEVNGVDYDS 200
DB 103 YKEYVTLCKETAEEAKGSDSLSPYELIEFNKRITW-----GQ-- 150
QY 201 RGQLEDEHTFEELKPLYEHLAAVVRRAKLNANAYSPYISPIGLCPAHLGLDMGKRWTL 260
DB 151 -----EHPYDALDLDFE--PGVYKVLVD-----QLPAEL 177
QY 261 YSLVYF-----GQRNIDVTDAVDQADQIRFEAKKEFVSVGLPNTQGFENS 313
DB 178 KEALIPVKQYVYASGNKP-----DTSFTTAFFPEKQKELSLYLOELGYDF-- 224
QY 314 MLTDPGNOKAVCHPTAMDLDGDRILMCTKVTMDDELTA-----HHEMGHIOYD--MAY 367
DB 225 ---DGRIDEYV--HFAITLNRGDYV--TRRYDEKDFRIALFETGHECGAITEQINDE 278
QY 368 AAQPLRLNGANGEGHEA-----VGE-----TMSLSAATPHKLSIGLSPD- 410
DB 279 ALSGINLSDGASMGHIESQSLFEYENFGRNKHFWPTPKKQKQKSPVQFNDISL--DDFV 336
QY 411 QEDNTEIEINFLKQK--LTYGILPFTYMLEKRWMTYKGEIPKQDMKKWKKREYV 468
DB 337 RAINESSKPSFIRVEADELTPPLHITIRYEIK--ALFSNVSVEDLPSLNMOKYODYL 393
QY 469 VVEPVPHDETCDPASLPHVSNDYS--FIRYRTTLQFOFOBALCO-----AAKHGRL 521
DB 394 I---TPQDASGITLQDVMAGSGDFYPSYALGYTAALQKQKIMEDLPEEDALLERGF 450
QY 522 HKCDISNSTAGOKLFNMLRL--GKSEPTLALENVGAKNNVRLPLNTEPL 569
DB 451 HPIK-----QWLEKVIHIGKRRKPLDIIKQATG--EELNVRLIDY 490

Search completed: May 26, 2003, 17:57:42
Job time: 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 16:53:03 ; Search time 21 Seconds

(without alignments)
1422.046 million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STIEEQAKTLDKFNHEAD.....DNSLEFLGIQPTLGPNNPP 720

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1337	34.6	732	1	ACET_HUMAN
2	1337	34.6	1306	1	ACE_HUMAN
3	1337	34.5	732	1	ACET_MOUSE
4	1334	34.5	1312	1	ACE_MOUSE
5	1312	33.9	1193	1	ACE_CHICK
6	1310	33.9	1313	1	ACE_RAT
7	1283	33.2	737	1	ACET_RABIT
8	1283	33.2	1310	1	ACE_RABIT
9	1073	27.7	615	1	ACE_DROME
10	1054.5	27.3	611	1	ACE_HAIE
11	135	3.5	501	1	YPMW_BACSU
12	135	3.2	627	1	GLGB_BACSU
13	121	3.1	1034	1	BGAL_BACME
14	119	3.1	986	1	EPIT_STAEF
15	115.5	3.0	3433	1	UTRO_HUMAN
16	115	3.0	950	1	Y511_RICPR
17	113.5	2.9	3655	1	YAMB_SCHPO
18	113	2.9	6669	1	NEBU_HUMAN
19	112	2.9	663	1	WTR_CABEL
20	112	2.9	3911	1	AKA9_HUMAN
21	111	2.9	4540	1	DYHC_PARTE
22	109	2.8	906	1	Y010_CLOAB
23	108.5	2.8	773	1	CND2_SCHPO
24	107.5	2.8	1398	1	DNM2_SCHPO
25	107.5	2.8	1689	1	RPV1_SCHPO
26	107	2.8	736	1	VP4_SCHPO
27	106.5	2.8	1284	1	CMFB_SCHPO
28	106.5	2.8	1084	1	XPOL_YEAST
29	105.5	2.7	1312	1	RA50_YEAST
30	105.5	2.7	645	1	REP_BUGAT
31	105	2.7	1002	1	POI_H1U4
32	105	2.7	1279	1	BCHH_CHLVI
33	105	2.7	2167	1	YCS2_YEAST

Not found

ALIGNMENTS

RESULT 1	ACET_HUMAN	STANDARD:	PRT:	732 AA.
ID	ACET_HUMAN			
AC	P22966;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kininase II).			
GN	DCPI OR DCP OR ACE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90046671; PubMed=2554286;			
RA	Ehlers M.R.W., Fox E.A., Strydom D.J., Riordan J.F.;			
RT	"Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=8938720; PubMed=2547653;			
RA	Lattion A.L., Soudrier F., Allegrini J., Hubert C., Corvol P.,			
RT	Alhenc-Gelas F.;			
RT	"The testicular transcript of the angiotensin I-converting enzyme encodes for the ancestral, non-duplicated form of the enzyme.";			
RT	FEBS Lett. 252:99-104(1989).			
RL	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.			
RX	MEDLINE=99251580; PubMed=10319862;			
RA	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;			
RT	"Sequence variation in the human angiotensin converting enzyme.";			
RT	Nat. Genet. 22:59-62(1999).			
RL	[4]			
RN	ZINC-BINDING.			
RX	MEDLINE=91308093; PubMed=1649623;			
RA	Ehlers M.R., Riordan J.F.;			
RT	"Angiotensin-converting enzyme: zinc- and inhibitor-binding stoichiometries of the somatic and testis isozymes.";			
RT	Biochemistry 30:7118-7126(1991).			
RL	[5]			
RP	FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF THE C-TERMINAL DIPEPTIDE.			
RX	THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN. ALSO ABLE TO INACTIVATE CATALYTIC ACTIVITY. Release of a C-terminal dipeptide, oligopeptide-1-xaa-xbb, when xaa is not pro, and xbb is neither Asp nor Glu. Converts angiotensin I to angiotensin II.			
CC	- CORFACTOR: BINDS 1 ZINC ION.			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and somatic (AC P12821); are produced by alternative splicing.			
CC	- TISSUE SPECIFICITY: SPERMATOCTES, ADULT TESTIS.			
CC	- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL			

	CC	-I- SIMILARITY:	BELONGS TO PEPTIDASE FAMILY M2.
	CC	-----	
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC	
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC	
	CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as their content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).	
	CC	or send an email to license@sib-sib.ch)	
	DR	EMBL; M26657; AAA60611.1;	--
	DR	EMBL; X16295; CAA34362.1;	--
	DR	EMBL; AF181656; AAD28561.1;	--
	DR	PIR; S05238; S05238.	
	DR	PIR; A33979; A33979.	
	DR	MEROPS; M02.004;	--
	DR	MIM; I06180;	--
	DR	IInterPro; IPRO01548; Peptidase_M2.	
	DR	IInterPro; IPRO00130; Zn_mpeptide.	
	DR	pfam; PF01401; Peptidase_M2_1.	
	DR	PRINTS; PR00791; PEPDIPASA.	
	DR	Prodrom; PD004184; Peptidase_M2_1.	
	DR	PROSITE; PS00142; ZINC_PROTEASZ; 1.	
	KW	Hypoxanthine aminotransferase; Carboxypeptidase; Zinc; Dipeptidase; Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing; Polymorphism.	
	FT	SIGNAL	1 31
	FT	CHAIN	32 732
	FT	DOMAIN	32 684
	FT	TRANSMEM	685 701
	FT	DOMAIN	702 732
	FT	METAL	414 414
	FT	ACT_SITE	415 415
	FT	METAL	418 418
	FT	ACT_SITE	442 442
	FT	CARBOHYD	457 457
	FT	CARBOHYD	103 103
	FT	CARBOHYD	121 121
	FT	CARBOHYD	140 140
	FT	CARBOHYD	186 186
	FT	CARBOHYD	368 368
	FT	CARBOHYD	617 617
	FT	VARIANT	32 32
	FT	VARIANT	49 49
	FT	VARIANT	712 712
	FT	SEQUENCE	732 AA; 83330 MW; 80E0D19CFA642313 CRC64;
OY	Query Match	Best Local Similarity	34.6%; Score 1337; DB 1; Length 732; Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9
DB	2 TIEBQAKPELDENHEDAEFDLYOSSLASNYVMNTEE-	NVONMNAAGDKWSA	53
OY	TIEBQAKPELDENHEDAEFDLYOSSLASNYVMNTEE-	NVONMNAAGDKWSA	53
DB	70 TDABAPSKVEERYDRTSQVNWEFYEAANNYNNTITTESKILLKKNOIAHHF--	LKNMAIHHT	123
OY	FKEOSTLAQMTPLOEIOTNLTVKLDAOLQQNGSSVLSEDDKSRRLLTIINTSTIYSTGK	II13	
DB	124 --LKVTQAARKPDVDNIQTWTIKRIKKOJDERALPAQELEEYVKRLIIDMETTYSVAT	181	
OY	VGNPNPQCCLLEPLGLANINMSUDYNERLTAWESMRSGEKOLRPLETEEVYLKNEHA	173	
DB	182 VCHPNG--SCLDLPDLIVNMALT SRKYEDLLAMWBGRKAGAILQFPKVVELLNQSA	239	
OY	RANHXYDYDGRCGYENVGVGGYYDSRGCLLEDVFHEETEELPKLEYEHAVVRATIMA	233	
DB	240 RLVNGYDVADDSDMKMSMYETSLE-----QLDELRRLQDELQPPLYLNLHAVRYRALHRH	289	
OY	234 Y-PSYISPGICCPAHLLGDMMKGREFTNLSLYVPFGQRKNIDVTAMDANDQADRIFE	292	

Db	290	YGAQHINLESPRIPIAHLLGNMMAQTSNIIYDLVPPSPASMDTTEAMLKOGWTPRRMFKE	349
Qy	293	AEEFYSVGLPNNTOGFEMSNULDPGNVOKAVCHPTAMDI-GK6--DRIILMCTRYIMDDF	351
Db	350	ADDEFFSLGLLPVPEPFNMKSMLEKPTDQREYVCHASMDYNGKDRIRIKOCTTVNEDL	409
Qy	352	LTAHEHGHITQYDMAAFAAPFLIRLNGANGEGHEAVGEIMSLSAATPKHLISIGLLSPDQ	411
Db	410	VVAHHEHGHITQYEMQKCDLPVALREGANGCFHEAIGDALVSLSTPKHLISHLNLSSEGC	469
Qy	412	EDNETENFLKCALITVGETLEPTVYLEKRMVNGKGIPLDQMMKKMMKKREIYVE	471
Db	470	SD-EHDINFLMKALCKIAIFIPSPSYVDQWRVRVDSITENTNQEWMSIRLYGSLCP	528
Qy	472	VPVHDEYVCPASLFFVSDNSYRIRYRTLLYQFOFQALCOAKHNEGPLHKDISNSTE	531
Db	529	VPVPTQDFFPGAKFHIISSVPPIREFVSIIQFFPHALCOAGHGPLHKDITYSKE	588
Qy	532	AGQKLFMLRLKGSSEPTLALENYVNGAKMMNVRPLNTFEPLFTWLKDQK--NSFVGM	588
Db	589	AGQALFAMKMLIGFSRPPMPPEAMOLITGQPMASAMLSYFKPLDMLRTENELHGEKLGWP	648
Qy	589	STDMSPYDQS 599	
Db	649	QYNTPTNSARS 659	

RESULT 2			
ACE_HUMAN			
Db	ID	ACE_HUMAN	STANDARD; PRT: 1306 AA.
Id	P12821;		
Dt	01-OCT-1989	(Rel. 12, Created)	
Dt	01-OCT-1989	(Rel. 12, Last sequence update)	
Dt	15-JUN-2002	(Rel. 41, Last annotation update)	
De	Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)		
De	(ACE) (Dipeptidyl carboxypeptidase I) (Kininase II) (CD143 antigen).		
Gn	DCPI OR DCP OR ACE.		
Os	Homo sapiens (Human).		
Oc	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Ox	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
Ox	NCBI_TaxID:9606;		
Rn	[1]		
Rp	SEQUENCE FROM N.A.		
Ra	MEDLINE:89071703;	Pubmed-2849100;	
Ra	Soubrier F., Alhenc-Gelas F., Hubert C., Allegri J., John M.,		
Ra	Tregear G., Corbol P.,		
Rt	"Two putative active centers in human angiotensin I-converting enzyme		
Rt	revealed by molecular cloning."		
Rt	Proc. Natl. Acad. Sci. U.S.A. 85:9386-9390(1988).		
Rn	[2]		
Rp	SEQUENCE FROM N.A. AND VARIANTS S-261; W-561 AND S-1286.		
Ra	MEDLINE:99251580;	Pubmed-10319862;	
Ra	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;		
Rt	"Sequence variation in the human angiotensin converting enzyme.";		
Rt	Nat. Genet. 22:59-62(1999).		
Rn	[3]		
Rp	PARTIAL SEQUENCE OF 30-46.		
Rc	TISSUE=Lung;		
Ra	MEDLINE:90110025;	Pubmed-2558109;	
Ra	Takachi K., Shimizu T., Ohishi N., Seyama Y., Takaku F.,		
Ra	Yotsumoto H.;		
Rt	"Purification of human lung angiotensin-converting enzyme by high-		
Rt	performance liquid chromatography: properties and N-terminal amino		
Rt	acid sequence."		
Rt	J. Biochem. 106:442-445(1989).		
Rn	[4]		
Rp	ZINC-BINDING.		
Ra	MEDLINE:91308093;	Pubmed-1649623;	
Ra	Elhers M.R., Riordan J.F.;		
Rt	"Angiotensin-converting enzyme: zinc- and inhibitor-binding		
Rt	stiochrometries of the somatic and testis isozymes."		
Rt	Biochemistry 30:7118-7126(1991).		

FT	VARIANT	561	561	R -> W.
FT				/Fid=VAR_011708.
FT	VARIANT	1286	1286	R -> S.
FT				/Fid=VAR_011709.
FT	CONFLICT	35	35	Q -> E (IN REF. 2).
FT	CONFLICT	42	42	D -> R (IN REF. 2).
SQ	SEQUENCE	1306 AA;	149714 MW;	IUB3BCAA301A26AA CRC64;

Query Match
Best Local Similarity 41.7%; Pred No. 2.7e-87; Length 1306;
Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9

OY	2	TTEBOAKTFLDKFNHAEEDLFYQSSLASWNTYNITEE-----NYQNMMNAGDKMSA	53
DB	644	IDEAESKEVEEYDRISQVWNNEYEANNNYNTNITETSKILLQKNQIAHT-----	697
OY	54	FLEKOSTLAOMPLDIOMLFYKLQLQALQONGSSVLSDRCKRLNTIINTMSTIYSIGK	113
DB	698	--LKTGTQARKRDVADOLQNTTKRIKKVODLERALLPAQOLEEYENKILIDMETTVSVAT	755
OY	114	VCNPDPPOECILLEBGLMEIMANSIDYNERRIAMESWSREYVKOLRPITYEEVYLKENMA	173
DB	756	VCHPMG--SCLDLPEDLFTVMATSKRYEDLLWAMSGMRDKAGRALLIQEPFKVELLIQNA	813
OY	174	RANHEDVDGYRGIVENGVDGYDSKQGLIEDVEHTPEETIKPLYEHIAAYRAKLMA	233
DB	814	RLNGVYDGSDSRSMYTEPPSLE-----QDLERLFOELQPLYLNHAVRRALHRH	863
OY	234	Y-PSYISPICLPAHLIDDMGRFNTLSLVTPFGOKPNIDVTAPMDADQIRIKE	292
DB	864	YGAGHINIEGPRIALLIGNMAQTNSITDYLVPPSPASMDTTAMLKOGWTPRMKRE	923
OY	235	AKEFVSVIGLPRMTOGFEWMSKLIDPGNVQKAVCHPTAMDLGK-DFFILMCTKYTMDF	351
DB	924	ADDEFTSLGLLPVPREFFNKSMLTEPRTGREGVCHASAMDFNGKDFRIKOTTYNLBDL	983
OY	352	LTAHEMHIOIDMAVYAOPFLIRMGANGDFEANGELMSLSAARPKHLKSGLLSPDQ	411
DB	964	VVAHHEMHIOTFMQYKDLPALRKGAHPGEFEALGDVALASVSPKHLHSINTLSSBG	1043
OY	412	EDNETEINFLLKQALTIVGLEPTYLMERKMVFKEGIPIKDWMKKMKREIYGVYE	471
DB	1044	SD-EHDINFELKMMALDKIAFIPEFSLYVDMMRWVRVDSITKENYQEWMSLRKTOGCLP	1102
OY	472	PVPHEVCDFPDASLFHVANDSYSETRPYRTILXQPOFALQOAAKHBERPLKCDISNSTE	531
DB	1103	PVPRKGQGFDEGAKFHIPSVPYITXFYSIIQEFHALCQAASHTPRLKRCDIYOJKE	1162
OY	532	AGOKLFNNLRGLKSEBPWTIALENVVGANKMNVRPLNFLEPLFTWLKDUNK--NSFWGM-	588
DB	1163	AGORLATAMKLGFSRPMPEAMQQLITGOGNMASAMLSTFRKLLDMLKRLENELHGKLCGW	1222
OY	589	STDMSPYADQS 599	
DB	1223	QYNMTPNARS 1233	

RESULT 3
ACET_MOUSE STANDARD: PRT: 732 AA.
ID ACET_MOUSE
AC p22967;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Angioensin-converting enzyme, testis-specific isoform precursor
DE (EC 3.4.15.1) (ACE-I) (Dipeptidyl carboxypeptidase I) (kininase II).
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
LN [1]
SEQUENCE FROM N.A.

RX MEDLINE-90318396; PubMed-2164636;
 RA Howard T.E., Shah S.-Y., Langford K.G., Martin B.M., Bernstein K.E.;
 RT "transcription of testicular angiotensin-converting enzyme (ACE) is
 RT initiated within the 12th intron of the somatic ACE gene.";
 RL Mol. Cell. Biol. 10:4294-4302(1990).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-I-Xaa-Xbb, when Xaa is not pro, and Xbb is neither
 CC asp nor glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 1 ZINC ION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and
 CC somatic (AC P09470); are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.
 CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL
 CC REGULATION BY ANDROGENS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M55333; AAA37149.1; -;
 CC EMBL; M61094; AAA37150.1; -;
 CC PIR; A35655; A35655.
 CC MEROPS; M02.004; -;
 CC MGD; MGI:87874; Ace.
 CC DR InterPro: IPR001548; Peptidase_M2.
 CC DR InterPro: IPR000130; Zn_Metpeptidase.
 CC Pfam: PF01401; Peptidase_M2; 1.
 CC DR PRINTS; PR00791; Peptidase_M2; 1.
 CC DR PRODOM; P004184; Peptidase_M2; 1.
 CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC DR Hydrolase; Metallopeptidase; Carboxypeptidase; Zinc; Dipeptidase;
 CC Glycoprotein; Transmembrane; Testis; Signal; Alternative splicing.
 CC KW SIGNAL 1 31
 CC FT CHAIN 32 732
 CC FT ANGIOTENSIN-CONVERTING ENZYME,
 CC TESTIS-SPECIFIC ISOFORM,
 CC EXTRACELLULAR (POTENTIAL).
 CC FT DOMAIN 32 684
 CC FT TRANSMEM 685 701
 CC FT DOMAIN 702 732
 CC FT METAL 413 413
 CC FT ACT SITE 414 414
 CC FT METAL 417 417
 CC FT METAL 441 441
 CC FT CARBOHYD 102 102
 CC FT CARBOHYD 120 120
 CC FT CARBOHYD 139 139
 CC FT CARBOHYD 185 185
 CC FT CARBOHYD 367 367
 CC FT CARBOHYD 616 616
 CC FT SEQUENCE 732 AA; 84047 MW; 16C817E7FBD09BD9 CRC64;
 SQ
 Query Match 34.5%; Score 1334; DB 1; Length 732;
 Best Local Similarity 42.6%; Pred. No. 2e-87;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;
 OY 2 TIEBAKTLKFKHEDLFYSSLSAMNTNTITEENQNNMADKMAFLKESTL 61
 DB 69 TDEAADRFEVEYDRTAQLVLLNERAEANWQNTITTEGSKILLEKSTEVSNHTLKGTR 128
 OY 62 AOMYPLQETOMLVYLOALQNGSSVLSSEKSKRLTTLTNTSTIYSGKVCNPNPQ 121
 DB 139 AKTFVSNFQNSIKRIKIKONDRALVLPKLEEVNQILLDETTYSLSNICYTG-- 186
 OY 122 ECLLEPGLNTEIMANSIDYNERLWMSWSEVQKQPLVEEYVVLAKEMARAHYEDY 181

DB 187 TCMPLRPLDTNMATSRKRYEELLMAMKSRDKVGALLPPFKYVERSNKIATKNGYTD 246
 OY 182 GDYMGDEYVNGVDGDYDSRGOLIDEVHTFEIEIKPLTEHLLAVYRAKIMNAPYS-YISP 240
 DB 247 GDSMSLYSESDLE-----QDEKLYQELQPLTLNHLAVYRSLHRYSEYINL 296
 OY 241 IGLCPAHLIGDMGFWTNLYSLVYFGQKPNIDVTAMDQANDAOIRFEAKRFYVS 300
 DB 297 DQPIPAHLIGNMMAQTGNSINLYDLVAPFSAPNIDTEAMIMQGTPRIFEAQNFETSL 356
 OY 301 GLPNTOGFWNSMLTDGNGVAKVCHPTADMLDGK--DEFLIMTKVMDQFLAHNMG 359
 DB 357 GLLPPEPEFWKSMLEKPLDGRVYCHSANDFYGKDFRKOCTSVAMQELVLAHHMG 416
 OY 360 HIOYMAVAAOPLFLRNGANEGFHEAVEISLSNATPKHLKISGLSPDFQEDNEIN 419
 DB 417 HIQYMQKDLDFVTFREGANPGFHEAIGDIMALSVSTPKHLISLNLSTE--GSGYEDIN 475
 OY 420 FLIKQALITVGLPTPTMLEKRWAVYFGELPKDQMKMKEMKRELYGVPEPHDET 479
 DB 476 FLMKALDKIARIPSYLIDQWRWRFVFGSTIKENYNOEMWSLRKLYGLCPVPRSGD 535
 OY 480 CPASLFHNSDYSTRYTRTYRILYQFOEALCQAKHEGPLHKCDISNSTEAGOKLFNM 539
 DB 536 FPDGSEFHPVAVPVYRVYSFIIOFQHEALCRAAGHTGPLHKCDIQSKAEAGKLADA 595
 OY 540 LRIGSEPTALLENVYAGKNNVRPLNYPEFLTWLKDQK--NSFVGW-STDWSP 594
 DB 596 MKLGYSKPWPEAMKILITGQPNMSASAMNVEFKPLEMTLVENRRRGELTGWPEYVWAP 653
 RESULT 4
 AC ACE_MOUSE STANDARD; PRT; 1312 AA.
 ID ACE_MOUSE
 AC P09470;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (dipeptidyl carboxypeptidase I) (kininase II).
 GN DCP1 OR ACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP MEDLINE-89308599; PubMed-2545691;
 RA Bernstein K.E., Martin B.M., Edwards A.S., Bernstein E.A.;
 RT "Mouse angiotensin-converting enzyme is a protein composed of two
 RT homologous domains.";
 RL J. Biol. Chem. 264:11945-11951(1989).
 RN [2]
 RP SEQUENCE OF 1-332 FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE-88298730; PubMed-2841312;
 RA Bernstein K., Martin B.M., Bernstein E.A., Linton J., Striker L.,
 RA Striker G.;
 RT "The isolation of angiotensin-converting enzyme cDNA.";
 RL J. Biol. Chem. 263:11021-11024(1988).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-I-Xaa-Xbb, when Xaa is not pro, and Xbb is neither
 CC asp nor glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
 CC specific (AC P22967); are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: J04946; AAA37147.1; -
 DR EMBL: J04947; AAA37148.1; -
 DR EMBL: J03940; AAA37146.1; -
 DR PIR: A29220; A29220.
 DR PIR: A34171; A34171.
 DR MEROPS: M02.001; -
 DR MEROPS: M02.004; -
 DR MGD: MGI:87874; Ace.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR PRINTS: PR00791; PEPIDPASEA.
 DR ProDom: PD004184; Peptidase_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 DR Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 34
 FT CHAIN 35 1312 ANGIOTENSIN-CONVERTING ENZYME, SOMATIC
 FT ISOPFORM.
 FT DOMAIN 35 1264 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1265 1281 POTENTIAL.
 FT DOMAIN 1282 1312 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 232 588
 FT REPEAT 830 1186
 FT METAL 335 395 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 336 396 1 (BY SIMILARITY).
 FT METAL 339 393 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 933 993 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 994 994 2 (BY SIMILARITY).
 FT METAL 997 997 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 59 59 ZINC 2 (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 682 682 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 700 700 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 719 719 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 765 765 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1196 1196 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 568 568 T -> A.
 SQ SEQUENCE 1312 AA; 150947 MW; 9C13BB0529AD3755 CRC64;

Query Match 34.5%; Score 1334; DB 1; Length 1312;
 Best Local Similarity 42.6%; Pred. No. 4,5e-87;
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

2 TTEEDAKTFLDKFNHEADLFTQSSLASMTNTNTEENVONMANNAGKWSAFLKEQSTL 61
 DB 649 TDEADAREVEYEDTAQVLLNEVFAEMQWNTNTTIESKLTLEKSTEVSHHTLKYGTR 708
 QY 62 AOMVPLQEIQLTVLQALQOQNGSSVLSBDSKRLMTLNTMTSTIYSGKVCNPDMPQ 121
 DB 709 AATFVSNFQNSIRIKTKLQNDRAVLPKELEEVQDILDMETYSLSNICTYNG-- 766
 QY 122 ECLLLEPGLEIMANSLDYNERIMAMESRSEVQKOLPLVEEYVVLAKENMARAHYDY 181
 DB 767 TCMPLPLEDLTNMAMATSRKYBELLMAMKSMROKVGAILPEPKYVEFNKAKLNGYIDA 826
 QY 182 GDYMGDYEVCVQDIDYDSKGLIDVETHEEIRPLEYHLHAYRAKILMANAYS-YISP 240
 DB 827 GDSMSRLESYDNLE-----ODLEKTLQEOLOPLYLNLHAYRSLHGHYGESEYINL 876
 QY 241 IGCPLAHLLGDMWGRFWNTLVSLVYVFGQKPNIDVTDMANVQAMQAQILFNEAEFFVS 300

DB 877 DGPIPAHLGMMWQOTNSNIYDLVAPPSAPNIDATAMKQGTPTPRIRKEADNFFSTL 936
 QY 301 GLPMPTOGFENSMILTPGNAVOKAHPITAMDLGK-DFFILNCTKTYMDFTAHHEMG 359
 DB 937 GLTPVPEFWMKSKLEKPTDGRREVCHPSANDFVNGDFRIKQCTSYNMEDLYIAHHEMG 936
 QY 360 HIOYDAMVAAOPELLRNGANGFEHVAEGLMSLSAATPKHLKSTIGLSPQSDNETEIN 419
 DB 997 HIOYFMQYKDLPTVFERGANGFEHVAIGDIALSVSTPKHLYSINLSTLSTLSTLSTL 1055
 QY 420 FLTKQALTYGTLPTFTYMLEKRWMPKGEIPKQNMKKMKREKRELVGVVEVPHDEY 479
 DB 1056 FLTKMALDKLAFPTFTSTLIDQWRWRVFDGSTITENYQEWMSLRLKYQGLCPPPSQGD 1115
 QY 480 CDPASLFHSNDYSFIRYETLYQFOFOEALCOANHEPRLHKDCISNSTEAGOKLFNM 539
 DB 1116 FDPGSKFVPANVYVYVFYSIIQFOFHALGCAACHTPTPKHCDIYQSKRKGKLADA 1175
 QY 540 LRLCKSEPTLALENVGAKMNVRLPLNTPLEPTWLKQDNK--NSFVGW-STDWSP 594
 DB 1176 MKLGSRKWPENAKLITGPQMSASAMNTEFKPTEWLVTENRHGFTLGMPVENAP 1233

RESULT 5
 ACE-CHICK STANDARD; PRT: 1193 AA.
 ID ACE-CHICK
 AC 010751;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl
 DE carboxypeptidase I) (Kininase II) (Fragment).
 GN DCP1 OR ACE.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID:9031;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RA MEDLINE-95110342; PubMed-7811282;
 RA Esther C.R., Thomas K.E., Bernstein K.E.;
 RT "Chicken lacks the testis specific isozyme of angiotensin converting
 RT enzyme found in mammals.";
 RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-1-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: I40175; AAA75554.1; -
 DR MEROPS: M02.001; -
 DR MEROPS: M02.004; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR ProDom: PD004184; Peptidase_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.

KW Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat.
 FT NON_TER 1
 FT DOMAIN <1 1156
 FT TRANSMEM 1157 1173
 FT DOMAIN 1174 1193
 FT REPEAT 125 481
 FT REPEAT 723 1079
 FT METAL 288 288
 FT METAL 289 289
 FT METAL 292 292
 FT METAL 886 886
 FT ACT_SITE 887 887
 FT METAL 890 890
 FT CARBOHYD 11 11
 FT CARBOHYD 60 60
 FT CARBOHYD 216 216
 FT CARBOHYD 407 407
 FT CARBOHYD 447 447
 FT CARBOHYD 485 485
 FT CARBOHYD 513 513
 FT CARBOHYD 555 555
 FT CARBOHYD 575 575
 FT CARBOHYD 658 658
 FT CARBOHYD 1089 1089
 SQ SEQUENCE 1193 AA; 137820 MW; 954472A1B8A471C7 CRC64;

Query Match 33.9%; Score 1312; DB 1; Length 1193;
 Best Local Similarity 40.4%; Pred. No. 1,4e-85;
 Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

OY 4 EEOAKTEPLDKFNEAEDELTYOSSLASNNVTNTTEENVOANNAGDKWSAFLEQSTLAQ 63
 DB 544 EAQAKTEPLDKFNEAEDELTYOSSLASNNVTNTTEENVOANNAGDKWSAFLEQSTLAQ 63
 OY 64 MYPLQELIONLVKQLQALQONGSSVLSDEKSKRLNLTMTSTVSGVCNPDN---P 120
 DB 604 QFDPSPDQDETIVYILKLKSLYERLALPEDELEKYNLTLSDMETYSVAACVCRBNTEFHP 663
 OY 121 QECLEELGELNEMANSIDYNERLAWESRSEVQKOLRPLYEYVLKKNEMARAHYED 180
 DB 664 -----LDPDLDIATSRDYNELLFKAKGMMWDSAGAKIKKRYVELSKAAVANGTYD 718
 OY 181 YGDYWRGDYEVNGVDYDSRGQILEDEVTFEIEKPLVEHLAAVYRAKIMAY-PSYIS 239
 DB 719 NGAYWRSLYETPTE-----EDLERLYLOPLLYLNLHAYRRALYNNKYGAEHIS 768
 OY 240 PIGGLPAHLIGDMGREFWNLVSLTFPGQKPNIDVDVAVDQAMDQRIFFKAEKFTVS 239
 DB 769 LKGPTRPAHLIGDMGREFWNLVSLTFPGQKPNIDVDVAVDQAMDQRIFFKAEKFTVS 239
 OY 300 VGLPNTQGFENSLMDPQNGVAKVCHPTAMD-L-GRGDFRILMCTKYVMDDEFLTAHHEM 358
 DB 829 LGLPMPQEFWDKSMIEKRPADGRVYCHASAMDPYNNKDRIRICQCTVYVNNDDLLTYHHEM 888
 OY 359 GHIOYDAVYAAQPLPLNKGANGEGHVAEGLMSISAAFPYHLKSGILSPDFQDNTEI 418
 DB 889 GHVYFLOQYMDQPSIFRDGANPGRHIGDYMALVSTPRHLHSINLD-QVTEENERSDI 947
 OY 419 NELLKQALTYVGLTPFYMLEKRMVYFKGEIPDDOMKRWKMEKREIVAVVPEPDER 478
 DB 948 NYLMSIALDKTAFLPEGLMDQMRKVFDSIKEDENQOQWMLNRLYQGLCPVPSPSED 1007
 OY 479 YCDPASLHVSNDYSTRYTRTLQYFOFQALCOAKHGGPILKDCISNSTFAGOKLFN 538
 DB 1008 DFDDGAKFHPANPYRIRYFVIOQFQALCKAGHGHPLTQDIYOSKRGKGLGD 1067
 OY 539 MLRLGKSEPTWLAENVGAKNNVRLVPEPLFWL--KQONKSFVGM-STDNSSPY 595
 DB 1068 AMKLGFSFPEPEAMQITIGOPNMSAEALMSFEPLMLVKKMTENGEVGLGWPEYSWT 1127
 OY 596 ADGSIKVRISLSALG-----DKAYEMNDENYLFTRSSVAVAMRQYELKVK 641

DB 1128 AVTEPHAVTDADFLGMSVGTNRQATAGAM----VLALALVELTISIFLGVR 1175
 RESULT 6
 ACE_RAT
 ID ACE_RAT STANDARD; PRT; 1313 AA.
 AC P47820;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (dipeptidyl carboxypeptidase I) (kininase II).
 GN DCP1 OR ACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=94121658; PubMed=8292044;
 RA Koike G., Krieger J.E., Jacob H.J., Mukoyama M., Pratt R.E.,
 RA Drazu V.J.;
 RT "Angiotensin converting enzyme and genetic hypertension: cloning of
 RT rat cDNAs and characterization of the enzyme."
 RL Biochem. Biophys. Res. Commun. 198;380-386(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEW/N; TISSUE=Lung;
 RA Jafarian-Tehrani M., Listwak S., Barrientos R.M., Michaud A.,
 RA Corvol P., Sternberg E.M.;
 RT "Characterization of a missense mutation in the angiotensin
 RT I-converting enzyme cDNA in exudative inflammation resistant F344/N
 RT rats."
 RL Submitted (NOV-1999) to the EMBL/Genbank/CDN databases.
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-N-xaa-xbb, when xaa is not pro, and xbb is neither
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
 CC specific; are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U03708; AAA82110.1; -
 DR EMBL: U03734; AAA82111.1; -
 DR EMBL: AF201332; AAG35597.1; -
 DR MEROPS: M02.004; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR PRINTS: PR00791; PEPTIDASEA.
 DR ProDom: PD004184; Peptidase_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 KW Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT SIGNAL 1 35
 FT CHAIN 36 1313
 FT DOMAIN 36 1265
 FT TRANSMEM 1266 1282
 FT POTENTIAL

FT	DOMAIN	1283	1313		CYTOPLASMIC (POTENTIAL).
FT	REPEAT	233	233	589	
FT	REPEAT	831	1187		
FT	METAL	396	396		
FT	ACT_SITE	397	397		
FT	METAL	400	400		
FT	METAL	994	994		
FT	ACT_SITE	995	995		
FT	METAL	998	998		
FT	CARBOHYD	44	44		
FT	CARBOHYD	60	60		
FT	CARBOHYD	80	80		
FT	CARBOHYD	117	117		
FT	CARBOHYD	152	152		
FT	CARBOHYD	166	166		
FT	CARBOHYD	324	324		
FT	CARBOHYD	515	515		
FT	CARBOHYD	683	683		
FT	CARBOHYD	701	701		
FT	CARBOHYD	720	720		
FT	CARBOHYD	766	766		
FT	CARBOHYD	948	948		
FT	CARBOHYD	1197	1197		
FT	VARIANT	207	207		
SEQ	SEQUENCE	1313	AA: 150907 MM: 8CB5D0015F129591 CRC64:		
Query Match		33.9%	Score 1310; DB 1; Length 1313;		
Best Local Similarity		42.0%;	Pred. No. 2.3e-85;		
Matches 251;	Conservative 115;	Mismatches 214;	Indels 18;	Gaps 7;	
2	TIEROAKTFIDKFNHEADLEFYOSSLASMNNTNTTEENVOANNAGCKMSAFLEKOSTL	61			
Db	550 TDEKKAANFEVEYRTAKVLAINEEAAMWANTNTTIGSKILLQKNREVSNHTLKGTW	709			
QY	62 AAMTPIAEIOMLTVKLOALQOONGSSVLSDEKSKRLTITLNTMSTIYSTGVCNPDMPQ	121			
Db	710 AKTDVSNFQOSTIKRIKKQVANDRAVLPRNELEBYQIILDMETTYISVANCTYNG--	767			
QY	122 ECLLEPGLNEMANSIDYNERNLWAMESWSEVGKOLRPLYEEVYVLNEMARANHYEDY	181			
Db	768 TCLSLPEPLTITMATSRYEELLTWKMSRQKRAILPFPKKYVDFSNKTLAKLNGYSDA	827			
QY	182 GDVNRGDYEVNGVDGYDUSROGLIEDVHTFEETKPLYLEHHAAYRATLMAAYP--YISP	240			
Db	828 GDSWSSSYSDLE-----QDEKLKVOELQPLVLNHAAYRRSLHRRYGSSEYNTL	877			
QY	241 IGCPLPAHLIGMMGFEWNTLXSLVPPFGOKFNIDVTDAMVOAMDADRIFKEAEKFEYSV	300			
Db	878 DGPRLPAHLIGMMMAQWNTNIIYDLVAPFSAISIDATEMIMIOGTPKRIFFEADDFETSL	937			
QY	301 GLPNMTOGFWENSMLTDGPNVOKAACHPTANDLGK--DFRIIMCTKYTMDDFLTAHHBMG	359			
Db	938 GLLPVPPEFMKSMLEKPTDGRREVYCHASAWDEYNGKDFRIKOSTSVMMELVLAHHBMG	997			
QY	360 HIQYDMAAAPPELLLRNANGFPHAVGEIMLSLATSATKHLKISGLSPDPOEDNETELN	419			
Db	998 HIQYFMOKRDLPVTFREGANGFHEAIDIVTALSVSTPKHLSHLSLSS--GSGYEHDIN	1056			
QY	420 FLIKOALITVGLTLPETYLEKRMWVFGFGEIPKDDKMMKWMEMKREIYGVVPEVPHDETY	479			
Db	1057 FLMKALDLKIAFIPPSYLIIDQRRMRVFGSGSTIKETKYNQEMMSLRKTYGIGLCPPVRSOGD	1116			
QY	480 CDPASLPHVSNDSYSEIRYRYRTLYQFOEALCAQAKKEGPLHKDCISNSTEAGOKLRPM	539			
Db	1117 FDPGSKFHVAPNVPIRYRIFESIIQFOHEALCRAAGHTGPLYKCDIYQSEAGKRLADA	1176			
QY	540 LRKSKSEPWTLALENVGAKMMNRPPLATREPLFTWKLKDNK--NSYVGW--SIDWSP	594			
Db	1177 MKIGYSKQWPMAPMAKKITIQPMMSASAINMYKRPLEMTLVTEENRRRGETLGGPEYTWIP	1234			

ID	ACCEP	RABIT	STANDARD:	PRN:	737 AA.
DT	P22968:				
DT	01-AUG-1991	(Rel. 19, Created)			
DT	01-AUG-1991	(Rel. 19, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor				
DE	(EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (kininase II).				
GN	DCPI OR ACE.				
OS	Oryctolagus cuniculus (Rabbit).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.				
OX	NCBI_Taxid:9986;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-New Zealand white; TISSUE-Testis;				
RX	MEDLINE-89380303; PubMed-2550457;				
RA	Kumar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;				
RT	"Structure of testicular angiotensin-converting enzyme. A segmental				
RT	mosaic isozyme."				
RL	J. Biol. Chem. 264:16754-16758(1989).				
RN	[2]				
RP	SEQUENCE OF 1-117 FROM N.A.				
RX	MEDLINE-91139683; PubMed-1847388;				
RA	Kumar R.S., Thekkumkara T.J., Sen G.C.;				
RT	"The mRNAs encoding the two angiotensin-converting isozymes are				
RT	transcribed from the same gene by a tissue-specific choice of				
RT	alternative transcription initiation sites."				
RL	J. Biol. Chem. 266:3854-3862(1991).				
CC	-1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF				
CC	THE TERMINAL HIS-RESIDUE, THIS RESULTS IN AN INCREASE OF THE				
CC	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.				
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,				
CC	oligopeptide- -xaa-xbb, when xaa is not pro, and xbb is neither				
CC	asp nor glu. Converts angiotensin I to angiotensin II.				
CC	-1- COFACTOR: BINDS 1 ZINC ION.				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS: 2 isoforms, testis-specific (shown here) and				
CC	somatic (AC P12822); are produced by alternative splicing.				
CC	-1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.				
CC	-1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL				
CC	REGULATION BY ANDROGENS.				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; J05041; AAA31153.1; -				
DR	EMBL; M58580; AAA31152.1; -				
DR	PIR; A34402; A34402.				
DR	MEROPS; M02.004; -				
DR	InterPro; IPR001548; Peptidase_M2.				
DR	InterPro; IPR001330; Zn_MPeptidase.				
DR	Pfam; PF01401; Peptidase_M2; 1.				
DR	PRINTS; PR00791; PEPDIPASFA.				
DR	ProDom; PD004184; Peptidase_M2; 1.				
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.				
KW	Glycolase; Metalloproteinase; Carboxypeptidase; zinc; Dipeptidase;				
KW	Hypocaprotein; Transmembrane; Testis; Signal; Alternative splicing.				
FT	SIGNAL	1	32		
FT	CHAIN	33	737		
FT					
FT	DOMAIN	33	690		
FT	TRANSMEM	691	707		
FT	DOMAIN	708	737		
FT	METAL	419	419		
FT	ACT SITE	420	420		
FT	METAL	423	423		
FT	METAL	447	447		
FT					

Query Match 33.28; Score 1283; DB 1; Length 737;
 Best Local Similarity 40.88; Pred. No. 8.5e-84;
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

FT CARBOHD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 737 AA; 83923 MM; FC43CC76655C3DCA CRC64;

2 TIEOAKTFLDKENHEADELFYOSLSAWNYNTNITE-----NVQMMNN--AGDKW 51
 75 TDEEASRFEVEYRSTQAVWNEYAEKMMNTNITEASKILLQNMQLANTHTLYGNW 134
 52 SAELEOSTLAOMYPLQIONTLVKLOLQALQNGSSVLSDESKRLNTILNTMSTYST 111
 135 -----ARRDVSNFQONATSKRIKKYODLQRAVLVYKLEEBYNOILLMETIYSV 184
 112 GVCNPNPDECELLLEGLEIMANSUDYNERLMWESRSENGKOLRPLYEERYVYAKNE 171
 185 ANVCRRVVG--SCLOLEEDLTNLTAKTSRKYDELMTWMTSMDDKGRALILPYEPYVEFTN 242
 172 MARANHEDYGDYRGDYEVNGVDYDYSRQOLIEDEYHTEFEETIKPLYEHLHAYVRKLM 231
 243 AARNGYVADGDSRSMETPTLE-----QDLERLPOLQPLYLMLHAYVRGALH 232
 232 NAY-PSYISPIGCLPAHLIDDMGRFTNLSLTPRGQRPNIDYDAMVADQADQRIE 290
 293 RHYGAHINLEGRIPAHILGMMNAQTWSNITDYLAPPSASTJDATEAMIKQGMTPRRMF 352
 291 KEAKFEVSVGLPMTGCFENSNLIDPGVQKAVCHPTAMDGKC-DEPIILCTKYTMD 349
 353 EEADEFISGLLEVPPEFNNKSMLEKPTGREVYASASAMDEYNGDFRIKCTTYNME 412
 350 DFLAHHEMGHIOYDMAAOPFLIRNGANEGFEHAGELMSLSAATPKLKSIGLSPD 409
 413 DLVYVHHEMGHIOYFMOYKDLPAALRGANPCHGHEALGDLTALSVSPKALHSINLISE 472
 410 FOEDNETEINFLKQALITVGLTFEYTMLEKRMWYFKEGEPIDQMKWMEKKREIYGV 469
 473 -GGGEHNDINFLMKMALDKIAIFPSYLVDEWHRVRVDSITKENVNOQWMSLRKYGL 531
 470 VEPHPHETCDPASLPHVSNDSFTYTRTLKXOPFOALCOAAKHESPLKCDISNS 529
 532 CPPAPRSQGDGDEPAKHFIPSSVYIXRYFSFIQFHALCKAGHTGPLETCDIYOS 591
 530 TEAGOKLFNNLRIGKSEPTLALENVYGAKNMVRPLNTEPEPLFTMLKDON--KNSFVG 587
 592 KEAKKRLADAMKLGYSKPEAKMVIITGOFNMSASAMANTFKPLMDLITENGRHGKLG 651
 588 W-STDSPPYADQS 599
 652 WQYTWTPNSARS 664

RESULT 8
 ACE_RABIT STANDARD; PRT; 1310 AA.
 ID ACE_RABIT
 AC P12822; 002852;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
 DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).
 GN DCP1 OR ACE.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;

RX MEDLINE=92178960; PubMed=1311831;
 RA Thekkumkara T.J., Livingston W. III, Kumar R.S., Sen G.C.;
 RT "Use of alternative polyadenylation sites for tissue-specific
 RT transcription of two angiotensin-converting enzyme mRNAs."
 RT Nucleic Acids Res. 20:683-687(1992).
 RN [2]
 RP REVISIONS.
 RA Sen G.C.;
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-88 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91139683; PubMed=1847388;
 RA Kumar R.S., Thekkumkara T.J., Sen G.C.;
 RT "The mRNAs encoding the two angiotensin-converting isozymes are
 RT transcribed from the same gene by a tissue-specific choice of
 RT alternative transcription initiation sites."
 RL J. Biol. Chem. 266:3854-3862(1991).
 RN [4]
 RP SEQUENCE OF 34-55.
 RC TISSUE=Lung;
 RX MEDLINE=84051289; PubMed=6314908;
 RA Iwata K., Blacher R., Soffer R.L., Lai C.Y.;
 RT "Rabbit pulmonary angiotensin-converting enzyme: the NH2-terminal
 RT fragment with enzymatic activity and its formation from the native
 RT enzyme by NH4OH treatment."
 RL Arch. Biochem. Biophys. 227:188-201(1983).
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
 CC THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
 CC oligopeptide-I-xaa-xbb, when xaa is not pro, and xbb is neither
 CC asp nor glu. Converts angiotensin I to angiotensin II.
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
 CC specific (AC P2268); are produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: X62551; CAA44428.1; -
 DR EMBL: M58579; AAA31151.1; ALT-SEQ.
 DR PIR: A23455; A23455.
 DR PIR: S35484; S35484.
 DR MEROPS: M02.001; -
 DR MEROPS: M02.004; -
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTpeptidase.
 DR Pfam: PF01401; Peptidase_M2; 2.
 DR PRINTS: PR00791; PEPDIPASEA.
 DR ProDom: PD004184; Peptidase_M2; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 2.
 KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.
 FT CHAIN 1 1310
 FT STAGL 1 33
 FT
 FT DOMAIN 34 1263
 FT TRANSMEM 1264 1280
 FT DOMAIN 1281 1310
 FT REPEAT 232 588
 FT REPEAT 829 1185
 FT METAL 395 395
 FT METAL 396 396
 FT ACT_SITE 399 399
 FT METAL 392 392
 FT METAL 992 992
 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 ZINC 1 (BY SIMILARITY).
 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 ZINC 2 (CATALYTIC) (BY SIMILARITY).

DE carboxypeptidase I) (Kininase II).
GN ANCE OR RACE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95293950; PubMed=7775412;
RA Cornell M.-J.; Williams T.A., Lamango N.S., Coates D., Corvol P.,
RA Soubrier F., Hohlisei J., Lehrach H., Isaac R.E.;
RT "Cloning and expression of an evolutionary conserved single-domain
RT angiotensin converting enzyme from Drosophila melanogaster.",
RL Biol. Chem. 270:13613-13619(1995).
[2]
RP REVISIONS.
RA Cornell M.-J.;
RP Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=Canlon-S;
RX MEDLINE=96028519; PubMed=7547664;
RA Tatei K., Gal H., IP Y.T., Levine M.;
RT "Race: a Drosophila homologue of the angiotensin converting enzyme.",
RL Mech. Dev. 51:157-168(1995).
CC -1- FUNCTION: MAY PLAY A ROLE IN THE CONTRACTIONS OF THE HEART, GUT
CC AND TESTES.
CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-1-xaa-xbb, when xaa is not pro, and xbb is neither
CC asp nor glu. Converts angiotensin I to angiotensin II.
CC -1- COPACITOR: BINDS 1 ZINC ION.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE AMNIOEOSA DURING GERM BAND
CC ELONGATION, SHORTENING AND HEART MORPHOGENESIS. EXPRESSED IN
CC MIDGUT THROUGHOUT EMBRYOGENESIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U25344; AAB02171.1; -;
CC EMBL; U34599; AAC46902.1; -;
CC MEROPS; M02_003; -;
CC DR FLYbase; FBPR0012037; Ance.
CC DR InterPro; IPR001548; Peptidase_M2.
CC DR InterPro; IPR00130; Zn_MTpeptidase.
CC DR Pfam; PF01401; Peptidase_M2; 1.
CC DR PRINTS; PR00791; PEPDIPPEASE.
CC DR ProDom; PD004184; Peptidase_M2; 1.
CC DR PROSITE; PS00143; ZINC_PROTEASE; 1.
CC KW Hydroxylase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
CC Glycoprotein; Signal.
CC FT SIGNAL 1 17
CC FT CHAIN 18 615
CC FT METAL 367 367
CC FT ACCT SITE 368 368
CC FT METAL 371 371
CC FT CARBOHYD 53 53
CC FT CARBOHYD 196 196
CC FT CARBOHYD 311 311
CC FT CONFLICT 48 51
CC FT CONFLICT 141 141
CC FT CONFLICT 293 293
CC FT CONFLICT 486 486
CC FT CONFLICT 533 533
CC FT CONFLICT 547 547
CC FT CONFLICT 615 AA; 71025 MM; 35830D5EA7F33CFB CRC64;
CC SEQUENCE

Query Match 27.7%; Score 1073; DB 1; Length 615;
 Best Local Similarity 36.5%; Pred. No. 6, 2e-69;
 Matches 218; Conservative 120; Mismatches 238; Indels 22; Gaps 9;

4 EQQAKFLDKFNEADLEFYQSSIASMNNTTEENVQNMNAGDKMSAFLEKQSTLQA 63
 22 EIQAKLEYENLEKLNKRTNVEEAMAGSNITDENKKNEISAELEKFKVEASDPT 81
 64 MYPLQETIONLVKLOLQAOQSSVLSDEKSKRLTITNTSTIYSTGVCPNDPQC 123
 82 KFWRSYQSEDLKRFKALTKLGYALPBDVAELDLDTLAMESSNAKVKVDYKSTGC 141
 124 -LLEPGLNEIMANSIDNRLMAMWSRSEVGKOLRPLYEEVYVVKENMARHEDYG 182
 142 DIALDPEIEVYSKSRDHEALYVWREFDKAGTAVRSOFERVELNTAAKLNNTSGA 201
 183 DYWRGDEVNGVDYDSRGQLIEDVEHFEETIKPLYEHLHAYVRKLMNAY-PSYISPI 241
 202 EAWLDEYE-----DDTEQQLEDI---FADIRPLYQOIHGVRFLRKHYGDVAVSET 251
 242 GCLPAHLGDMGRFTNLSLVPRGOKPNIDVTAMVDQAMDARIKREAEKFPVSG 301
 252 GFLPMHLGMMMAQMSIADIVSPPEKPLVDVSAEMEKQATTPKMFQMGDFPTSMN 311
 302 LPMNTOGFEMENSLTPGNAVQKAVCHPTAMDLGK-GDFRIIMCTKYTMDFTLAHMGH 360
 312 LTKLPQDFMKSITIEKPTDGRDLVCHASAMDFLLIDVRLKQCTRTYQQLFVHHELG 371
 361 IYDMAYMAOPFLIRNGANGFHEAVGEIWSLSAATPKHLKSGLSPPQEDNEETINF 420
 372 IYFELYQHPFVYKRGANGFHEAVGDVLSVSPKHLKELGLK-DYVRDEARINQ 430
 421 LKMAQITVOTLPTVTLKRWMAVFGSEIPIKDOMKKKWMEMREIVGVPEPHDETC 480
 431 LEFLALDKIYFLPFAFTMDKTYRMSLRGEVDKANNMCATPKLDEYSGIEPPVRSKDF 490
 481 DPASLFVNSDYFIRYRTTYLYQFOEALC-OAAKH-----EGPLHKDISNSTEAG 534
 491 DAPAKYHISADVELRYLVSIIIOFQYKACIKAGQYDPELVNLPDNDIYGASARA 550
 535 KLRMLRLGSEPTLALENVGAKNNRPLNTEPEPLTWLAKDN--KNSVVGST 590
 551 AFHNMLSMGASKWPDPALAEFNGERLMSKALAEIPEPLKVLLEAKNNVHIGWT 608

RESULT 10

ACE_HAETE STANDARD: PRT: 611 AA.
 AC Q10715;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Angiotensin-converting enzyme precursor (EC 3.4.15.1) (Dipeptidyl
 carboxypeptidase I) (Kinase II).
 GN ACE.
 OS Haemaphysa irritans exigua (Buffalo fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscoidae; Muscidae; Muscidae; Haematobia.
 OC NCBI_TaxID=34678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215437; PubMed=8647080;
 RA Wajfelds G.L., Fitzgerald C., Gough J., Riding G.A., Elvin C.,
 Kemp D.J., Willadsen P.;
 RT Cloning and characterization of angiotensin-converting enzyme from
 the dipteran species, Haematobia irritans exigua, and its expression
 in the maturing male reproductive system.
 RL Eur. J. Biochem. 237:414-423(1996).
 CC -1- FUNCTION: INVOLVED IN THE SPECIFIC MATURATION OR DEGRADATION OF A
 NUMBER OF BIOACTIVE PEPTIDES.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,

oligopeptide-1-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
 Asp nor Glu. Converts angiotensin I to angiotensin II.
 -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 -1- TISSUE SPECIFICITY: EXPRESSED IN THE COMPOUND GANGLION AND IN THE
 POSTERIOR REGION OF THE MIDGUT.
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

EMBL: L43965; AAA70427.1; -.
 DR MEROPI; M02.003;
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_Mpeptidase.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS; PR00791; PEPDIPASA.
 DR ProDom: PD004184; Peptidase_M2; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
 KW Glycoprotein; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 611
 FT METAL 367 367
 FT ACT_SITE 368 368
 FT METAL 371 371
 FT CARBOHD 53 53
 FT CAROHD 156 196
 FT CAROHD 531 531
 SQ SEQUENCE 611 AA; 70505 MW; A43D6DF5A83ECB53 CNG64;

Query Match 27.3%; Score 1054.5; DB 1; Length 611;
 Best Local Similarity 37.7%; Pred. No. 1, 3e-67;
 Matches 222; Conservative 105; Mismatches 239; Indels 23; Gaps 8;

1 STEBOAT-FLDKFNEADLEFYQSSIASMNNTTEENVQNMNAGDKMSAFLEKQ 59
 18 ARKEIVATEYVIONNKELAKHTNVEEAMAGSNITDENKKNEISAELEKFKVEA 77
 60 TLAQMPLOEIONLVKLOLQAOQSSVLSDEKSKRLTITNTSTIYSTGVCPNDP 119
 78 KDIOKFNKTYGSAVVRQFSLKSTGYSALEADYALTELVLSAMESNFPKAVRCYKN 137
 120 PQC-LLEPGLNEIMANSIDNRLMAMWSRSEVGKOLRPLYEEVYVVKENMARH 178
 138 SAKCDLSIDPEIEEITTSRDEELKTYWTFYKAGTPTSNEPKYELNTRKAKLNF 197
 179 EDYGDYWRGDEVNGVDYDSRGQLIED-VEHFEETIKPLYEHLHAYVRKLMNAY-PS 236
 198 TGAEWLDEYE-----DATFEQQLAEIREDIKPLYDQHGVRFLRKHYGDVA 246
 237 YISPIGCLPAHLGDMGRFTNLSLVPRGOKPNIDVTAMVDQAMDARIKREAEK 296
 247 VVSKGFLPMLGMMMAQMSIADIVSPPEKPLVDVSAEMEKQATTPKMFQMGDF 306
 297 FVSVGLPNTQGFEMENSLTPGNAVQKAVCHPTAMDLGK-GDFRIIMCTKYTMDFT 355
 307 FQSMGLKRLPQFEMKSLIEKPTDGRDLVCHASAMDFLLIDVRLKQCTRTYQQLF 366
 356 HEMGHIQDMAYMAOPFLIRNGANGFHEAVGEIWSLSAATPKHLKSGLSPPQEDNE 415
 367 HEMGHIQYPLQYOHOPFVYKRGANGFHEAVGDVLSVSPKHLKELGLK-NVYSNE 425
 416 TEINFLKALITVGLPTVTLKRWMAVFGSEIPIKDOMKKKWMEMREIVGVPEPH 475
 426 ARNOLFETALDKIYFLPFAFTMDYRWALFQGDSEKMCATPKLDEYSGIEPPV 485
 476 DETYCDPASFVNSDYFIRYRTTYLYQFOEALCOAA-----KHEGDLHKDISNS 529

DB 486 TEKEDPAKHYASADVEYLRIVLSFIIQFOFYKACINAGEVVPNOTETPLNCDIYGS 545

OY 530 TEAGOKLFNNMLRLKSGSEPTLALENVVGAKNMVRPLNTPEPLTWL 578

DB 546 KEAGLFFNNMLSLGASKRFPDPALEAFNGERTMGKALAEYEPDLRWLE 594

RESULT 11

ID YPWA_BACSU STANDARD; PRT; 501 AA.

AC P50848;

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical metalloprotease yPWA (EC 3.4.24.-).

GN YPWA.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168 / Marburg;

RX MEDLINE=96349105; PubMed=8760912;

RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,

RA Serror P.;

RT "Sequence analysis of the Bacillus subtilis chromosome region between

RT the sera and kag loci cloned in a yeast artificial chromosome.";

RL Microbiology 142:2005-2016(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Berttero M.G., Bessieres P., Boloetin A.M., Borchert S.,

RA Boursier U., Brans A., Brun M., Brignell S.C., Bron S.,

RA Brouillet S., Brusch C.V., Caldwell I.F., Cummings N.J., Daniel R.A.,

RA Choi S.K., Codani J.J., Connerion I.F., Ehrlich S.D., Emerson F.T.,

RA Denzot F., Devine K.M., Dusterhoft A., Ferrati E., Foulger D.,

RA Entian K.D., Errington J., Fabret C., Ferrari E.J., Galleron N.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizki A., Grandi G.,

RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,

RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasaahara Y., Klier-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krog S., Kumano M.,

RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Meisel D., Nakai S., Noback M.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Paro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Roche B., Rose M., Sadleir F.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,

RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,

RA Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A.,

RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

RA Winlers P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus

RT subtilis.";

RT Nature 390:249-256(1997).

CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL: L47838; AAB38482.1; -

DR EMBL: L77246; AAB96610.1; -

DR EMBL: Z99115; CAB14125.1; -

DR MEROPS: M32.0PW; -

DR Subtilisin: B01.1458; YPWA.

DR InterPro: IPR001333; Peptidase_M32.

DR InterPro: IPR001330; Zn_MTPeptide.

DR Pfam: PF02074; Peptidase_M32; 1.

DR PRINTS: PR00998; CR00XPTASSET.

DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.

DR Hypothetical protein; Hydrolase; Metalloprotease; Zinc;

KW Complete proteome.

KW METAL.

FT ACT_SITE 265 265 ZINC (CATALYTIC) (POTENTIAL).

FT METAL 266 266 ZINC (CATALYTIC) (POTENTIAL).

FT METAL 269 269 ZINC (CATALYTIC) (POTENTIAL).

SO SEQUENCE 501 AA; 58174 MW; A/489B8BFA38F82 CAC64;

Query Match 3.5%; Score 135; DB 1; Length 501;

Best Local Similarity 20.9%; Pred. No. 0.036;

Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

OY 162 YEEHYVL--KNEMA--RANHEDYD-----DYRGDYEYNGVDGYDS 200

DB 103 YREYVILCSKAEFAWEAKGSDSFLSPYLEQLIEFKREITW-----GIG-- 150

OY 201 RGLIDVHTFEETKPLEYEHAVYRAKIMAYPSYISPGCLPAHLIGDMGRFTNL 260

DB 151 -----BHPYDALDLE--PGYVAVLD-----QLPAEL 177

OY 261 YSLVTFP-----GQPNIDVTDAVDQAMDQRIEFAKEFYVSGLPNNMGFWENS 313

DB 178 KEAIIPLVQVYVTSNGKP-----DTSFTKAFPEKOKELSLYELQELGYDE-- 224

OY 314 MLTPDGVQKAVCHPAPNMLGKQDFRIIMCTVTMDPFLTA-----HHMGHIQYD--MAY 367

DB 225 ---DGGRLDLYT--HPFATLNRGDVY--TRIDEKDFATLFGTHIEGHAITYQNIDE 278

OY 368 AAOFLFRNANGCFEHA-----VGE-----INSLAATPKHLKSTIGLSDF-- 410

DB 279 ALSTNTSLSDASNGIHESQSLFENFLGNKHHPWPKYKIDQASPVQFKDSTL--DDFV 336

OY 411 QEDNETINFLKQA--LITVGLPFTYMLEKRMVFGELPKQOMKMKMKMKREITV 468

DB 337 RAINEKSPRIYEADELTPYLIIRYELER--AIFSNESVDELPGLMOKYQDYL 393

OY 469 VVEVPVPHDEYCDPASLFVSNDS--FIRYRTLYQFOFOEALCO-----AAKHEGL 521

DB 394 I---TPQTDAGLIDYVHAGSDGFGTPSYALGMYAAQLKQMLDELPFALLERGEF 450

OY 451 HPIK-----QWLTEKVHIGKRRKRPDIIRDATG--EELNVRLIDY 490

RESULT 12

GLGB_BACSU STANDARD; PRT; 627 AA.

AC P39118;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.16) (Glycogen branching

DE enzyme).

GN GLGB.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=94195107; PubMed=8145641;

RA Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;

RT "Glycogen in Bacillus subtilis: molecular characterization of an

```

RT RT operon encoding enzymes involved in glycogen biosynthesis and
RT degradation";
RL RL Mol. Microbiol. 11:203-218(1994).
RN RN [2]
RP RP SEQUENCE FROM N.A.
RX RX STRAIN-168;
RC RC MEDLINE=98048467; PubMed=9387221;
RA RA Lapidus A., Galleron N., Sorokin A., Ehlich S.D.;
RT RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rnb-dnaB region";
RL RL Microbiology 143:3431-3441(1997).
RN RN [3]
RP RP SEQUENCE FROM N.A.
RX RX STRAIN-168;
RC RC MEDLINE=98044033; PubMed=9384377;
RA RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Bouras L., Brans A., Brann M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denhart D., Devine K.J., Dunthorn A., Ehrlich S.D., Emmerson P.T.,
RA Entlin K.D., Erington J., Fabret C., Ferrati E., Foulger D.,
RA Feltz C., Fujita M., Fujita Y., Fuma S., Galliz E.J., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Goultly E.J., Grandi G.,
RA Giuseppe H., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert E., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karmata D., Kashara Y., Klaerr-Blandhard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsen G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarovic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwilk S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivaola C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Sekiguchi Y., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
RA Sorokin A., Taccon E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpestra P., Tognoni K.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
RA Vlati A., Wambit R., Wedler E., Wedler H., Wetzinger T.,
RA Winters P., Wipert A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshitawa H., Zanchin A.;
RT RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL RL Nature 390:249-256(1997).
RN RN [-]
RP RP FUNCTION: CATALYZES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC
RC RC LINKAGES IN GLYCOGEN BY SCISSION OF A 1,4-ALPHA-LINKED
RA RA OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE
RT RT SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF
RL RL ALPHA-1,4-GLUCAN CHAINS.
RN RN [-]
RP RP CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of
RC RC glycogen.
RA RA [-]
RT RT PATHWAY: Glycogen biosynthesis; third step.
RL RL [-]
RN RN INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON
RC RC SOURCES THAT ALLOW EFFICIENT SPOULATION.
RA RA [-]
RT RT SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
RL RL KNOWN AS THE ALPHA-AMYLASE FAMILY.
RN RN [-]
RP RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RC RC between the Swiss Institute of Bioinformatics and the EMBL outstation
RA RA at the European Bioinformatics Institute. There are no restrictions on
RT RT use by non-profit institutions as long as its content is in no way
RL RL modified and this statement is not removed. Usage by and for commercial
RN RN entities requires a license agreement (See http://www.isb-sdb.ch/announce/
RC RC or send an email to license@sdb-sib.ch).
RA RA -----
RX RX EMBL; Z57595; CAA81040.1;
RC RC EMBL; AF008820; AAC00214.1;
RA RA EMBL; Z99119; CAB15076.1;
RT RT PIR; S36624; S36624.
RL RL Subtilist; BG10907; g19p.
RN RN InterPro; IPR000461; Alpha_amlase

```

DR	InterPro: IPR004193; Isomylase.N.
DR	Pfam: PF00126; alpha-amylase; 1.
DR	Pfam: PF02922; Isoamylase_N; 1.
KW	Glycogen biosynthesis; Transferase; Glycosyltransferase;
KM	Complete proteome.
FT	ACT_SITE 309 309 BY SIMILARITY.
FT	ACT_SITE 352 352 BY SIMILARITY.
FT	ACT_SITE 420 420 BY SIMILARITY.
SQ	SEQUENCE 627 AA; 73665 MW; 64B0A553B6767BA CRC64;
Query Match	Score 125; DB 1; Length 627;
Best Local Similarity	18.6%; Pred. No. 0.25;
Matches 131;	Conservative 86; Mismatches 225; Indels 264; Gaps 37
QY	31 NNNNITTEENVQNNNAAGDKWSAF-----KEOSTLAOMYPLOETIONLTVKIQLALAOON 85
DB	56 DFNSSGSEEHVMEHVNDNGIMTLEIFIGIGEKER-----YKYEIVTN-----N 97
QY	86 GSSVLISBDKSKRINTLNTMTSTIYSFGKVCPDNPQECILLBPGLNEIMAN--SLDYNER 143
DB	98 GEIRLRKADP-----YAIYS-----EVAHPNLSLYDLE 125
QY	144 LMAMESKRSSVGQOLRPLEEYVYLKNEMARANHEDYGDYRGDYEAVGVGYDYSRGQ 203
DB	126 GYSMODOKWKOKAKATLEYKPFVI-----YELHGSMK-----KHDGRRTSYKE 171
QY	204 LIDYV-----EHFEE--EIPLTYEHHAIVRAKLMAAPSYISPICLPAILLDGMGRF 256
DB	172 LSQTILPIYIKKHGFHTIELLPVE--HPYDS-----WGQY 205
QY	257 WYNLYSLTPFGCKPNIDVTDAVDOA-----WDAGRIFEAKEKFVSGLPMMT 306
DB	206 GGEGYSPTSRFG--PPHDLM-KFYDECHQONGIVLLDWVRGHFCDAHGLYMGDEP--- 259
QY	307 QGFWEWSMLTDGQNVQKANCFPTANDLKGFDRITMCTKYTMDFLAN--HEMGHIQ 362
DB	260 --LYEKERRDEN-----WLWGTANFDL---CKPVHSFLISNALYMAEFYHI- 303
QY	363 YDMAVAOPEFLRLRANGNEGFI-EAVGEIMLSIAATPKHLKSIGLSDPOEDNETEIFNL 421
DB	304 -----DGFRVDVANILTMPNDERH-----INRYAVDFL 333
QY	422 LKQALITVGTLPETFMNL--EKRRKRVFGEIIPKD-----QMKKWMEMREIVGVEPY 473
DB	334 KKLNTMTAEAYPHVMMLIEDSTEMPVOYGAVEEGGLGFHKMNKGW--ANDVLKMYEP 390
QY	474 PHDETCDPASLFHYNSDYSFIRYTRTLVQFOEQALCQAARKHGGFLHNCDSNSTENG 533
DB	391 PEERRHC-----HQLSFSLL-----YASEFPVLPFSHDEVY-----G 425
QY	534 OKLFNMALLGSESPRTLALENVCAKANNVRPLLNY-----FEPLFTMK 578
DB	426 KK--SLTKMKMDGW-----OKPAQYRLLLGUYTWVHRPKKLIJFMGSFENQFDEM-K 473
QY	579 DONKSNFNGNTDSPYADOSI-----KVRSLKSALGDAYTEMND--NEYML 624
DB	474 DTEQ--LDWELDSFRMHQKASVTFODLLRFYOKRSKILVENHDRASFEMIDVHNDEOSI 530
QY	635 FRSSVAYAMROYFLKVNOMILFGEEDYRVANKLPRISEFFETAP 670
DB	531 F-SPIRYGOKH-----GEALVVICNFTPVVYHHQIDYGVP 563
RESULT 13	
ID	BGAL_BACME STANDARD; PRT; 1034 AA.
AC	052847;
DT	15-DEC-1998 (Rel. 37, Created)
DT	15-DEC-1998 (Rel. 37, Last sequence update)
DE	15-DEC-1998 (Rel. 37, Last annotation update)
GN	Beta-galactosidase (EC 3.2.1.23) (lactase).
GN	Bacillus megaterium.

CC Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC NCBI_TaxId=1404;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-DSM 319;
 CC STREY J.;
 CC Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC galactose residues in beta-D-galactosides.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ000733; CAA04267.1; -
 CC HSP: P00722; 18GL.
 CC InterPro: IPR004200; Bgal_small_C.
 CC InterPro: IPR004199; Bgal_small_N.
 CC InterPro: IPR001649; GH_2.
 CC Pfam: PF00703; Glyco_hydro_2; 1.
 CC Pfam: PF02836; Glyco_hydro_2_C; 1.
 CC Pfam: PF02837; Glyco_hydro_2_N; 1.
 CC Pfam: PF02929; Bgal_small_N; 1.
 CC Pfam: PF02930; Bgal_small_C; 1.
 CC PRINTS: PR00132; GLHYDRLASE2.
 CC PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
 CC PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
 CC HydroLase: Glycosidase.
 CC ACCT_SITE 481 481 PROTON DONOR (BY SIMILARITY).
 CC ACT_SITE 547 547 NUCEOPHILE (BY SIMILARITY).
 CC FT SITE 1034 AA; 118673 MW; 38644C9A69415E9 CRC64;
 CC SO SEQUENCE

Query Match 3.1%; Score 121; DB 1; Length 1034;
 Best Local Similarity 19.2%; Pred. No. 0.99;
 Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;

OY 11 LDKRHEHEDFYQSSSL-----ASWYNNINTEENQNNN---NAGDKSAFLKE 57
 DB 260 LNYEHDTQDLTFEMLDANAQEVLPLOTINSSQRTVSJLTHIKSPAKKSA----- 315
 OY 58 OSTLAQWYPLQEIQNL-TVKLQALQONGSSVLSSEDKSK--RLNTLLNTMSTIYSTGK 113
 DB 316 -----ESPRLTYLVL---SLKNAAGSIITETSCKVGPFREELKNGILMTI--NGK 359
 OY 114 -----VCPNDNPOECILLEPGLNEIMANSIDYERIM----- 145
 DB 360 RIVLRGNRHEPDSYKGRAGITREDMIDILMKOHNNINAVRTSHYPMDSWYELCNEYG 419
 OY 146 -----AM-----ESMRSEYVKOLRPLYEE----- 164
 DB 420 LYYIDETNLHGTWYTYLOEGEQKAVPGSKPKKENVLDRCRSMKERKNPISIIWISLG 479
 OY 165 -----YVVLK-NEMARANHYEDYGDYMGDYEVNGVDGYDYSRGOLLDEVHT 211
 DB 480 NESFGGENQMYTFEKEDSTRLVHYE--GIFHHRDYDASDIESTMYKP---ADVE-- 532
 OY 212 FEELKPLYLEHLIAYRAKALMAVPSYISIGCLPRLHGLGDMGRWTNLYSLTVPFGOKP 271
 DB 533 -----RYALMNPKRKYTL---CEYSHAMGNSCG---NLYKTYWEILFDOTP 570
 OY 272 NID---VTDAMVQDAQRIKFEAE--KEFVSVG-----LEPMTQGEWENSLTDPGNY 321
 DB 571 ILQGGRTMD-----WKDQALQATADGDSYLAAGDFDGTNP-DGNFCNGGLIFADGTA 623
 OY 322 OKAV-----CH-PFLAN---DLKGDRIILMCTKRYTMDLTLAHHEGHL-----QYDAY 367
 DB 624 SPRIAEVKKCYGCVVKTAVDPAPAKGFAY-----OKRHLEFTNLINAYDFW 667

OY 368 AAQPLRLNGANGNEHVEAGVINSLSATPKHKLSTIGLSPPROEDNETEINFLKQALT 427
 DB 668 TVE-----KNG-----ELVERKASLILNVPAGCTDELTSLTYLQENBID-EFVLTLSLR 716
 OY 428 I 428
 DB 717 L 717

RESULT 14
 EPIB, STAP
 ID EPIB, STAP STANDARD; PRT: 986 AA.
 AC P30195;
 DT 01-APR-1993 (rel. 25, Created)
 DT 01-APR-1993 (rel. 25, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Epidermin biosynthesis protein epib.
 GN EPIB.
 OS Staphylococcus epidermidis.
 OG Plasmid pTU 32.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 CC NCBI_TaxId=1282;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-TU 3298 / DSM 3095;
 CC MEDLINE-92155237; PubMed-1740156;
 CC Schell N., Engleke G., Augustin J., Rosenstein R., Ungermann V.,
 CC Goetz F., Britian K.-D.;
 CC "Analysis of genes involved in the biosynthesis of lantibiotic
 CC epidermin";
 CC Eur. J. Biochem. 204:57-68(1992).
 CC -1- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
 CC LANTIBIOTIC EPIDERMIN.
 CC -1- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
 CC THE CYTOPLASMIC SIDE OF THE MEMBRANE.
 CC -1- SIMILARITY: TO B. SUBTILIS SPAB AND L. LACTIS NISB.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X62386; CAA44253.1; -
 CC Plasmid.
 CC KW SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;
 CC SO SEQUENCE

Query Match 3.1%; Score 119; DB 1; Length 986;
 Best Local Similarity 18.2%; Pred. No. 1.3; 247; Indels 310; Gaps 40;
 Matches 150; Conservative 115; Mismatches 115; Indels 310; Gaps 40;

OY 2 TIEQANTFLDKENH-----AEPL--FYQSSLSAWYNNINTEENQNNNAGDKK 51
 DB 339 TIRNYHEFFMDKYFEDLVNKLKLLSDINGEGYPRKXSYFSNNI----- 383
 OY 52 SAKLEOSTLAQWYPLQEIQNLTVKLOALQONGSSVLSSEDKSKRLNTLLNTMSTIYST 111
 DB 384 -AFLEK-----YLLAIONSHLEITENDYKNLEK-NNTYVSKNA- 421
 OY 112 GKVCPNDNPOECILLEPGLNEIMANSIDYNERLMAWESWSENGKOLRPLYEEVYLKNE 171
 DB 422 -----PVSTEL-----YSRIY-----GNSIKG-YEDFAVVISPI 449
 OY 172 MARANHYEDYGDYMGDYEVNGVDGYDYSRGOLLDEVHTFEELKPLYEHLIAYRAKLM 231
 DB 450 LGSFVAGATFGF--TGNFNK-----KKNQLQREIYH-----HNNYMNENGL 491
 OY 232 NAYPSYISIGCLPRLHGLGDMGRWTNLYSLTVPFGOKPNIYTDAMVQDAQRIK 291
 DB 492 ELSQLEPGLNSRNINILNN--NRILYNCMLNLP---KSDIDINDIFIGATFVKLYLS 546

FT REPEAT 2443 2556 SPECTRIN 17.
 FT REPEAT 2559 2636 SPECTRIN 18.
 FT REPEAT 2658 2688 SPECTRIN 19.
 FT REPEAT 2691 2797 SPECTRIN 20.
 FT DOMAIN 2812 2845 WM.
 FT ZN_FING 3064 3111 22-TYPE.
 SQ SEQUENCE 3433 AA; 394488 MW; EAE8DBA09F8585B5 CRC64;
 Query Match 3.0%; Score 115.5; DB 1; Length 3433;
 Best Local Similarity 18.5%; Pred. No. 13; Indels 265; Gaps 36;
 Matches 139; Conservative 98; Mismatches 249;
 11 LKFNHEADL-----FYQSSLASWNTNTTEENVOANNAGDKWSAFLEKQSTL 61
 2533 LDMNORNDMDAKKASISRAHLEASAEKRNRLMSLEIKLANKDEE---LKKOMPT 2588
 62 AQMYPIQETONTLVKIQALQALQNGSSVSEDKSKRLNTILMTSTIYSGKVCNPDNPQ 121
 2589 GGDVPAIQLQYDHCALRRELKEKEYSVLNAVDAQRY-----FLADQPIEAPEPR 2639
 122 ECLLIEPGL-----NEIMANSIDYNERLWAMES-----WRSEVQQLRPYEE 164
 2640 RNLQSKTELTPERRAQKAKAMKRSSEVEK---WESINAVTSNOKQVDALKEKLRL 2696
 165 YVVLKNEARANNHEDYGYWR--GDYEVNGVGYDYSNGQLIEVHEFEELKPLYEHL 222
 2697 QGAMDDLDADKKEASVSNWGMKPVGDLIDSLQDH-----LEKIMAREELAPL 2745
 223 HAYYRAKLNNAYPSYISPIGLP---AHLIDMGHFWNTNLSLTVPGQKFNIDVYA 278
 2746 --NRKVTYNDLSQSLSPIDHPSLKMSRQDLIDLMNR-W-KLQVSV-----DD 2790
 279 WVDQAMDQRIKFEAREKFEVSVGL-----PNNTOGFV 310
 2791 RLKQLOEARNDFEGSSQHFLSTSVOLPMQNSISHNKPYIINHOYTQTCDDHFKMELF 2849
 311 ESMMLTDPGNV-----OKAVC-----HFTANDLKGDF 338
 2850 --QSLADLNNVFSAYRTAIRRLQKALCLDLELSTNEIFKOH---KLQNDQLLS 2903
 339 --RIIMCTRVMTDFLTAHHEMGLQ-----YDMVAAQDFELRANGANGCFH 383
 2904 VPDVINCILTYTDGLEOMHKDLVNPCLVCMKMLNAVYDTG-----RTKIRYQS 2955
 384 EANGELISLS-----AATPKHL---KSIGLSPD----- 409
 2956 LKIG-LMSLSKGLLEEKYRLEKFAVAGPTMCDQRLGILHDAIOIPROLEVAAGGS 3014
 410 -----FOEDNETEINFLKQALTYGILPFTYMLEKRWAVKGEIPADQMKRW 460
 3015 NIEPSVRSQFOQNN-----KPEISVKERI-DMM 3042
 461 EMKREIVGVEPVPHDETCDPA---SLPHVSNDSFYRYRTILYQFOFQALQAAKH 517
 3043 HLEQDSH-VMLPVLHRAAEETAKHQAQKICEKCEIYGFRRYSKLHFYND--VQOS--- 3096
 518 EGPIAKCDISNSTAGOKL-FNNMLRL-----GKS-EPWTALENVVAGAKNNMVR-PLLN 568
 3097 -----CFEFGSRTPAKGKRLHYPWVEYCIPTTSGEDVADPLKYVLKNNFRSKYFAKPRIG 3150
 569 YFEPLFTMLKQDNKNSFVGSSTDSVPADDS 599
 3151 YL-PVQYVLEGDNIETPTLLISMWPEHYPDS 3180
 RESULT 16
 Y511_RICPR STANDARD; PRT; 950 AA.
 AC Q92D36;
 DT 30-MAY-2000 (rel. 39, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Hypothetical protein Rp511.

GN Rp511.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiaceae; Rickettsia.
 OC NCBI_Taxid=782;
 RX STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E.; Zomorodipour A.; Andersson J.O.;
 RA Sticheritz-Ponten T.; Alsmark U.C.M.; Podowski R.M.; Naeslund A.K.;
 RA Eriksson A.-S.; Winkler H.H.; Kurland C.G.;
 RT "the genome sequence of Rickettsia prowazekii and the origin of
 mitochondria.";
 RL Nature 396:133-140(1998).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AJ235272; CA14963.1;
 DR InterPro; IPR001646; Speptide.repeat.
 DR Pfam; PF00805; Pentapeptide.1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 950 AA; 108612 MW; C047F8BCF063F715 CRC64;
 Query Match 3.0%; Score 115; DB 1; Length 950;
 Best Local Similarity 18.6%; Pred. No. 2.3; Indels 268; Gaps 37;
 Matches 145; Conservative 108; Mismatches 258;
 4 EEOAKTEFLDKRNHE-----AEDLFYQ--SSLASWNTNTTEENVOANNAGDKWSAFLEK 56
 120 EGGLEIKTECFRNEVYTKLQDIALYKQKEDTNNKYS---NILDMLTKRNFQKFN 175
 57 EOSTLAQMYPIQETONTLVKIQALQALQNGSSVSEDKSKRLNTILMTSTIYSGKVCN 116
 176 EKSQKQITNH-----IKSGATELLPSDYMSKSPFDIILQKPKKKLKLKTFNT---H 220
 117 PDNPOECL--LLEPGLNEIMANSIDYNERLWAMESRSEVQQLRPYEEYVVLKNNMAR 174
 221 PDVKQELVANNINNPILK-KENKLEPYKQELM-1SLKEVAQAQSPFLOE----- 268
 175 ANHEDYGYWRGDYEVNGVGYDYSNGQLIEVHEFEELKPLYEHLAAVYRAKLNNAY 234
 269 --HFESY-----KIDKILDIPTLLNKIPIKEIFDTLNA----- 302
 235 PSYISPIGLCPAHLIDMGHFWNTNLSLTVPGQKPN---IDVTDAMVDAQMDQRIE 290
 303 -----PKNGVMSISLEKALEWVAGDDQ--- 323
 291 KEAKFEVVS--VGLPNNTOGFWENS-----MLTDPGNV--OKAVCHPRAW 331
 324 --LKSFPANKKTLIPNNALGIIENFTPSVOSTNEVNFQDMVLYVEGWSKFEIHEIITA 381
 332 DLKGGDFRILMCTVY-----MDDFLTAHHEM-----GHIOYDM-AVAAQPL 373
 382 DLNKGDYMSLTGNIISITINDPSFKLADILYQSKGLFDNLGVLEQDIKNSQILKQDL 441
 374 LRGANGEGFHEAVGEIMSLSAAPPKHLKSIGLSPPFOEDNETEINFLKQALTYGILP 433
 442 INYGMENGDVYRKLKIMPIILDKDESILKY--FRDPIKGNVY----- 481
 434 FTYMLEKRWAVKGEIPKQDMKMKWMEKREIVGVEPVPHDETCDPASLPHVSNDS 493
 482 -----KMKRELISLTKNDKPKIKRYELN-----NNRAI 507
 494 FIRYTRTYLQFOFQALQAAKHGBPLAKCDISNSTAGOKLFNNMLRLKQSEPTVL-A 551
 508 FASIIDKTLMDLP-----GINNLDK--OELYNILPSMLNHPDELITY 567

QY 552 LENYV-----GAKN--MANVRPLNTEPEPLFTMLKQONKNSF-VGSHSTMSPYADQSIKRV 603
 Db 548 IEEVEKSHYHGAVASAIYNLAOKNTYE-----GOLPITIKAGNSGFN-YATEKVDV 599
 QY 604 ISLSALGDKAYEMNDNEMNTLFRSSVAYAROFPLKVKNMILFEEDVVRANKPKRIS 662
 Db 600 FSSSRDPKDKVID-----ETIRKHLKIQN-----GKFLBAILLGNLSN 641
 QY 663 FNEFYAPKNVSDIIPREVEKAIKMSRSRIDARLNDNSLEFLGID-----PTLGP 715
 Db 642 IDFSGVSLKN-ADLFVTSI-KDCNFKNTNLVDA-KLPDNLIMETDTYTNLDKAIPLTAP 697

RESULT 17

YAMB_SCHPO STANDARD: PRT: 3655 AA.

AC Q10064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C1F5.11c in chromosome 1.
 GN SPAC1F5.11c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens S., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymopre B.,
 RA Wellens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Muehlbauer S.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leclercq V., Mottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas R., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrall B.G., Nurse P.,
 RA The genome sequence of Schizosaccharomyces pombe.
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: STRONG. TO YEAST YHR099M.
 CC -1- SIMILARITY: BELONGS TO THE P13/P14-KINASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: Z68136; CA92239.1;
 DR InterPro: IPR003151; FAT.
 DR InterPro: IPR003152; FATC.
 DR InterPro: IPR000403; P13_P14_kinase.

DR InterPro: IPR001440; TPR.
 DR Pfam: PF00454; P13_P14_kinase; 1.
 DR Pfam: PF02259; FAT; 1.
 DR Pfam: PF02260; FATC; 1.
 DR SMART: SM00146; PI3K; 1.
 DR PROSITE: PS02900; P13_4_KINASE_3; 1.
 DR Hypothetical protein: Transferase, kinase.
 FT DOMAIN 3324 3655
 SO SEQUENCE 3655 AA; 420774 MW; 50475E3F3C2124A CRC64;

Query Match 2.9%; Score 113.5; DB 1; Length 3655;
 Best Local Similarity 16.9%; Pred. No. 20;
 Matches 90; Conservative 92; Mismatches 193; Indels 157; Gaps 22;

QY 2 TIEEQAKT---FLDKRNEHAEDELFT---QSSIASWNTNTITEENVOYNNMAGDKKSAF 54
 Db 2641 TLEKATKSLSPFSLRRHTADALYLNTKQRMKGSVTEFSRIIDECQFSIR---RQOL 2697
 QY 55 LKE--OSTLAOMYPLQEIQNLTVKQLQALQONGSVLSSEKSKRLNTIINTWSTYSG 112
 Db 2698 PKRYQSHVSLHHQEIYEL-----QEARGISQINDTNNINIDNKL----- 2740
 QY 113 KVCNPNDPQCLLEPGLNEIMANSLDYNERLMA-WESWSEYQKOLRPLEEYVLKNE 171
 Db 2741 -----RDIVVQLQGRERLPNWVD-DIDWEDLWAMQSVKSKNVFLPLVSIQAO 2791
 QY 172 MARANHYEDYGDYMGDEYVNGVDGYDSRQGLDEVEHTFEIKPLYLEHLAAYRAKIM 231
 Db 2792 STNKSNTSVSYLYNGIHEL-----AWII 2815
 QY 232 NAYPYSIPICGLPAHLIGDMGRFWNTLYSLTVPEGOKPNIDVTAMVDQMDAQRFK 291
 Db 2816 NNF-AHAVRHLEPEVCINDL---TKIVTL-----PNEIQCAFLEKDEQECYHE 2862
 QY 292 EAEKFFVSGVPPNNT-----QGFWSNLTDPGNVQAVCPAPAMDKG 335
 Db 2863 SPSEWQGLLEVINTNTNLMYRNRQKAEFFLKGMPFN-RLGKRDANOAFATAVQIDIGS 2921
 QY 336 GDFRLMCTKYMDLFTANHEKGIQYDMAYVAOPFLIRNAGN---EGFHEVGEINSLS 393
 Db 2922 G-----KAWSEWG-LYHDELFOANQOEIHNCNVCSEFLQ-SSLSSS 2963
 QY 394 AATPKHLKSLGLSPDEQEDNEFEINFLKQALITVGLPTFYMLEKRWVFGGEIPKD 453
 Db 2964 NSKFLRLRVMLLSVDSHGSSEV-----VSSFSKSELPWNMIPF---IPQ- 3007
 QY 454 QMMKMKEMKRELYVEVPEHDETCYCPASLFHVSNDY-SFIRYRTTYQ 504
 Db 3008 -----LLSALSHRESIHARAAILIQAKTYPSLHPQLRTAYE 3044

RESULT 18

NEBU_HUMAN

ID NEBU_HUMAN STANDARD: PRT: 6669 AA.

AC P20929; 015346;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nebulin.
 GN NEB.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=95257391; PubMed=7739042;
 RA Labat S., Kolmerer B.;
 RT "The complete primary structure of human nebulin and its correlation
 RL to muscle structure".
 RN J. Mol. Biol. 248:308-315(1995).
 [2]

RP PARTIAL PRELIMINARY SEQUENCE.

RA MEDLINE-88284704; PubMed-397062;
RA Zentani M., Darras B.T., Rizzuto R., Salvati G., Betto R.,
RA Bonilla E., Miranda A.F., Du J., Samitt C., Dickson G., Walsh F.S.,
RA DiMauro S., Francke U., Schon E.A.;
RT "Cloning and expression of human nebulin cDNAs and assignment of the
RT gene to chromosome 2q31-q32.";
RL Genomics 2:249-256(1988).
RN [3]
RP STRUCTURE BY NMR OF 6610-6669.
RX MEDLINE-98179539; PubMed-9514727;
RA Pollitt A.S., Millevoi S., Gautel M., Kolmerer B., Pastore A.;
RT "SH3 in muscles: solution structure of the SH3 domain from nebulin.";
RL J. Mol. Biol. 276:189-202(1998).
CC -1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MAINTAINING
CC THE STRUCTURAL INTEGRITY OF SARCOMERES AND THE MEMBRANE SYSTEM
CC ASSOCIATED WITH THE MYOFIBRILS. BIND AND STABILIZE F-ACTIN.
CC -1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC. LOCATED IN THE THIN FILAMENT
CC OF STRIPPED MUSCLE.
CC -1- DISEASE: Defects in NEB are a cause of the autosomal recessive
CC form of nemaline myopathy (NEM1).
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 178 NEBULIN REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X83957; CAA58788.1; -;
DR EMBL; M19668; AAA59916.1; ALT_SEQ.
DR EMBL; M19669; AAA59917.1; ALT_SEQ.
DR PIR; A29979; A29979.
DR PIR; B29979; B29979.
DR PDB; IARK; 28-JAN-98.
DR PDB; INEB; 24-DEC-97.
DR GeneW; HGNC:7720; NEB.
DR MIM; 161650; -;
DR MIM; 256030; -;
DR InterPro; IPR000900; Nebulin.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00880; Nebulin; 146.
DR PRINTS; PR00510; NEBULIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00227; NEBU; 181.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50002; SH3; 1.
KW Actin-binding; Muscle protein; Cytoskeleton; Repeat; 3D-structure.
FT REPEAT 76 107
FT REPEAT 113 143
FT REPEAT 148 178
FT REPEAT 183 213
FT REPEAT 218 248
FT REPEAT 253 283
FT REPEAT 289 318
FT REPEAT 324 354
FT REPEAT 363 393
FT REPEAT 398 428
FT REPEAT 434 464
FT REPEAT 502 532
FT REPEAT 537 567
FT REPEAT 573 603
FT REPEAT 611 641
FT REPEAT 681 711
FT REPEAT 749 779
FT REPEAT 784 814
FT REPEAT 820 850
FT REPEAT 858 888
FT REPEAT 893 923
FT REPEAT 924 954
NEBULIN 22.
NEBULIN 2.
NEBULIN 3.
NEBULIN 4.
NEBULIN 5.
NEBULIN 6.
NEBULIN 7.
NEBULIN 8.
NEBULIN 9.
NEBULIN 10.
NEBULIN 11.
NEBULIN 12.
NEBULIN 13.
NEBULIN 14.
NEBULIN 15.
NEBULIN 16.
NEBULIN 17.
NEBULIN 18.
NEBULIN 19.
NEBULIN 20.
NEBULIN 21.
NEBULIN 22.

FT REPEAT 959 990
FT REPEAT 993 1023
FT REPEAT 1028 1058
FT REPEAT 1064 1094
FT REPEAT 1102 1132
FT REPEAT 1137 1167
FT REPEAT 1168 1198
FT REPEAT 1204 1234
FT REPEAT 1237 1267
FT REPEAT 1272 1302
FT REPEAT 1308 1338
FT REPEAT 1346 1376
FT REPEAT 1381 1411
FT REPEAT 1412 1442
FT REPEAT 1448 1478
FT REPEAT 1481 1511
FT REPEAT 1516 1546
FT REPEAT 1552 1582
FT REPEAT 1590 1620
FT REPEAT 1625 1655
FT REPEAT 1656 1686
FT REPEAT 1692 1722
FT REPEAT 1725 1755
FT REPEAT 1760 1790
FT REPEAT 1796 1826
FT REPEAT 1834 1864
FT REPEAT 1869 1899
FT REPEAT 1900 1930
FT REPEAT 1936 1966
FT REPEAT 1969 1999
FT REPEAT 2004 2034
FT REPEAT 2040 2070
FT REPEAT 2078 2108
FT REPEAT 2113 2143
FT REPEAT 2144 2174
FT REPEAT 2180 2210
FT REPEAT 2213 2243
FT REPEAT 2248 2278
FT REPEAT 2284 2314
FT REPEAT 2322 2352
FT REPEAT 2357 2387
FT REPEAT 2388 2418
FT REPEAT 2423 2453
FT REPEAT 2456 2486
FT REPEAT 2491 2521
FT REPEAT 2527 2557
FT REPEAT 2565 2595
FT REPEAT 2600 2630
FT REPEAT 2631 2661
FT REPEAT 2666 2696
FT REPEAT 2699 2729
FT REPEAT 2734 2764
FT REPEAT 2770 2800
FT REPEAT 2808 2838
FT REPEAT 2843 2873
FT REPEAT 2874 2904
FT REPEAT 2909 2939
FT REPEAT 2942 2972
FT REPEAT 2977 3007
FT REPEAT 3013 3043
FT REPEAT 3051 3081
FT REPEAT 3086 3116
FT REPEAT 3117 3147
FT REPEAT 3152 3182
FT REPEAT 3185 3215
FT REPEAT 3220 3250
FT REPEAT 3256 3286
FT REPEAT 3294 3324
FT REPEAT 3329 3359
FT REPEAT 3360 3390
FT REPEAT 3395 3425
FT REPEAT 3428 3458
FT REPEAT 3463 3493
NEBULIN 23.
NEBULIN 24.
NEBULIN 25.
NEBULIN 26.
NEBULIN 27.
NEBULIN 28.
NEBULIN 29.
NEBULIN 30.
NEBULIN 31.
NEBULIN 32.
NEBULIN 33.
NEBULIN 34.
NEBULIN 35.
NEBULIN 36.
NEBULIN 37.
NEBULIN 38.
NEBULIN 39.
NEBULIN 40.
NEBULIN 41.
NEBULIN 42.
NEBULIN 43.
NEBULIN 44.
NEBULIN 45.
NEBULIN 46.
NEBULIN 47.
NEBULIN 48.
NEBULIN 49.
NEBULIN 50.
NEBULIN 51.
NEBULIN 52.
NEBULIN 53.
NEBULIN 54.
NEBULIN 55.
NEBULIN 56.
NEBULIN 57.
NEBULIN 58.
NEBULIN 59.
NEBULIN 60.
NEBULIN 61.
NEBULIN 62.
NEBULIN 63.
NEBULIN 64.
NEBULIN 65.
NEBULIN 66.
NEBULIN 67.
NEBULIN 68.
NEBULIN 69.
NEBULIN 70.
NEBULIN 71.
NEBULIN 72.
NEBULIN 73.
NEBULIN 74.
NEBULIN 75.
NEBULIN 76.
NEBULIN 77.
NEBULIN 78.
NEBULIN 79.
NEBULIN 80.
NEBULIN 81.
NEBULIN 82.
NEBULIN 83.
NEBULIN 84.
NEBULIN 85.
NEBULIN 86.
NEBULIN 87.
NEBULIN 88.
NEBULIN 89.
NEBULIN 90.
NEBULIN 91.
NEBULIN 92.
NEBULIN 93.
NEBULIN 94.
NEBULIN 95.

```

FT REPEAT 3499 3529 NEBULIN 96.
FT REPEAT 3537 3567 NEBULIN 97.
FT REPEAT 3572 3602 NEBULIN 98.
FT REPEAT 3603 3633 NEBULIN 99.
FT REPEAT 3638 3668 NEBULIN 100.
FT REPEAT 3671 3701 NEBULIN 101.
FT REPEAT 3706 3736 NEBULIN 102.
FT REPEAT 3742 3772 NEBULIN 103.
FT REPEAT 3780 3810 NEBULIN 104.
FT REPEAT 3815 3845 NEBULIN 105.
FT REPEAT 3846 3876 NEBULIN 106.
FT REPEAT 3914 3944 NEBULIN 107.
FT REPEAT 3949 3979 NEBULIN 108.
FT REPEAT 3984 4014 NEBULIN 109.
FT REPEAT 4021 4052 NEBULIN 110.
FT REPEAT 4057 4087 NEBULIN 111.
FT REPEAT 4088 4118 NEBULIN 112.
FT REPEAT 4123 4153 NEBULIN 113.
FT REPEAT 4156 4186 NEBULIN 114.
FT REPEAT 4191 4220 NEBULIN 115.
FT REPEAT 4226 4256 NEBULIN 116.
FT REPEAT 4294 4329 NEBULIN 117.
FT REPEAT 4330 4360 NEBULIN 118.
FT REPEAT 4365 4395 NEBULIN 119.
FT REPEAT 4400 4430 NEBULIN 120.
FT REPEAT 4435 4465 NEBULIN 121.
FT REPEAT 4471 4501 NEBULIN 122.
FT REPEAT 4544 4574 NEBULIN 123.
FT REPEAT 4575 4605 NEBULIN 124.
FT REPEAT 4610 4640 NEBULIN 125.
FT REPEAT 4645 4675 NEBULIN 126.
FT REPEAT 4680 4710 NEBULIN 127.
FT REPEAT 4716 4746 NEBULIN 128.

```

Query Match 2.9%; Score 113; DB 1; Length 6669;
 Best Local Similarity 18.4%; Pred. No. 51;
 Matches 169; Conservative 123; Mismatches 261; Indels 364; Gaps 50;

```

OY 50 KMSAFLEQSTLAQMTPLQEIOMLTV-KTOLQMLQONGSSVLSVE--DKSKRLNTLTNTM 105
DB 2387 QMWC-LPDNDVYQAKKVELESENLYKSDLEWLRGIGSPGSLAEKKNR-----A 2438
OY 106 STYISGKYCNPDPNOCILLERGLAEIMA--NSLDYNERLM--AMESRSEYKQOLRL 161
DB 2439 SEIISSEKKYKQPPDRKKFTSIPDAMDIVLAKTKAKRSRLRYKREAWDKKQTOI--HIMPD 2496
OY 162 YEEYVVLKKNEMARANYEDYGDWRYGVNGVDGYD-----Y 199
DB 2497 TPRIY-----LAKANLINTSDKLYRMGYEELKRRKGYDLVDALPIKAKASRIASEYKY 2551
OY 200 SRG--QL-----IED-----VEHTEERKPLYEH-----LH 223
DB 2552 KEGFRKQOLGHIGIARNIEDDPKMMMSMHYAKIQSDREYKKDFEKWKTKSSPYDMGLVYL 2611
OY 224 AYVRAKIMN--APRYSISPGICPAHLHGMGRFNTNLSLVPRPGQPNIDVPM--V 280
DB 2612 AYKCOQLTVSDVCKNKLHOMTCLP-----DOSVYIHA 2643
OY 281 DQAMDAQ--RIKFEAEKFEVSVGLPMTQGFWMN-----SMITDPGNVOKAVCHP--- 328
DB 2644 RQAVDLOSDMLYKSDLOMLKGIOW--WTSGSLEDENKKNRATQTLSD--HYVRQ--HPDGF 2697
OY 329 -----TAMDLAGKGFRIIMCT-----KVMDDPL 352
DB 2698 KFSIAMDIPMVLAKNNAITMNRHLYTEAWDKDKTYVIMPDPREVLAKONKRVANSEKI 2757
OY 353 TANHMEGHIO--YMAYAAQPELLRNGA-----NEGPHNAVGEIMSLSA--T 396
DB 2758 ---YKIGLEEKRRKGYMDALPIKAKASRIASEFKYKSGIKQOLGHIGIARNLRD 2814
OY 397 PKHLSIGLSPDQEDNETEINF-----LKQALTIYGLTPFTYMLE 439

```

```

DB 2815 PKMMMSMHYAK--IQSDREYKKDFEKWKTKSSPYDMGLVYLAKKCOQLTVSDVYKNNYLH 2872
OY 440 KWRM-----VFKGEIIPKQOMMK--WM-----EMKR--EIVG 468
DB 2873 QMTCLPOOSDVYIARQAYDLOSDMYKSDL--QMRHGIGWSIGSLDVEKCKRAEILIS 2929
OY 469 VVEPPEDEYCDASLF-----HVSND 491
DB 2930 -----DKIYRQPPDRFFKFTSVDSLEQVLAKNNALNNKRLYTEAWDKKQTOIIMP 2982
OY 492 YSFIRY-----YTRLLXQFOFQALCOQAKHEG-----PLHKDISNSTEAGOKL 536
DB 2983 TPETMLARONKINSEFLYKLANE-----AKKGGDLRSDAIPYAAKASRQVYSYKY 3037
OY 537 FNMRLKSEPTLALENVGAKNNVRLINTEPFLTWLKDQ--KNSEFVGSTWSPY 595
DB 3038 -----KQGYRQOLGHIGIARNIEDDPKMMMSMHYAKIQSDREYKKDFEKWKTKSSP 3089
OY 596 AD-----QSIKVRISLKSALGDKATEW-----NNEMYLFPSSVAYAMROYFLK 639
DB 3090 VDMGLVYLAKKCOQLTVSDVYKNNYL--HEWTCPLPDQ-----YIARQAYDLO 3136
OY 640 VKN-----QMTLFGEDY-----RVANLKPRISEFVYAPKNVSDIIPREVE 683
DB 3137 SDNIYKSDLOMLRGIGWPIGSMOYVCKRAAEI--LSDNITROP-----DKLKTFSYV 3189
OY 684 KAIRMSRSRINDAFRLN 700
DB 3190 DSLQGVYLA--NNALNMN 3205

```

RESULT 19

MTNR CAEEL STANDARD; PRT; 663 AA.

AC Q17693;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

PT Probable methyltransferase hydrofolate reductase (EC 1.5.1.20).

GN C0648.1.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditiidae;

OC Rhabditiidae; Pelodermidae; Caenorhabditis.

OX NCBI_Taxid=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Leimbach D.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP REVISIONS.

RA Waterston R.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADPH(+)

CC -1- COFACTOR: FAD (BY SIMILARITY).

CC -1- PATHWAY: Folate metabolism.

CC -1- SIMILARITY: BELONGS TO THE METHYLENETHETRAHYDROFOLATE REDUCTASE

CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb.ch/announce/>

CC or send an email to license@isb.ch).

CC -----

CC EMBL: U39849; AAA81048.2; -

CC HSSP: P00394; 1B5T.

CC Wormpep: C0648.1; C030593.

CC Interpro: IPR004621; Fadh2_euk.

CC Interpro: IPR003171; Methylrof_redctse.

CC -1- ALTERNATIVE PRODUCTS: 6 ISOFORMS; 1 (SHOWN HERE), 2, 3/CG-
NAP, 4/YOTIAO, 5 AND 6/ARAP350; ARE PRODUCED BY ALTERNATIVE
SPLICING.
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. ISOFORM 4/YOTIAO IS HIGHLY
EXPRESSED IN SKELETAL MUSCLE AND IN PANCREAS.
CC -1- DOMAIN: RII BINDING SITE. PREDICTED TO FORM AN AMPHIPHATIC HELIX.
CC -1- COULD PARTICIPATE IN PROTEIN-PROTEIN INTERACTIONS WITH A
COMPLEMENTARY SURFACE ON THE R-SUBUNIT DIMER.
CC -1- CAUTION: REF.6 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO TWO
FRAMESHIFTS IN POSITIONS 3782 AND 3811.
CC -1- CAUTION: REF.9 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FOUR
FRAMESHIFTS IN POSITIONS 29, 1653, 1659 AND 1735.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ131693; CAB40713.1; -
DR EMBL: AB019691; BAA78718.1; -
DR EMBL: AJ010770; CA009361.1; -
DR EMBL: AF026245; AAB86384.1; -
DR EMBL: AF083037; AAD22767.1; -
DR EMBL: AF004013; AAB96867.1; - ALT_FRAME.
DR EMBL: AF091711; AAD39719.1; -
DR EMBL: AB018346; BAA34523.1; -
DR EMBL: AC000066; AAC60380.1; - ALT_FRAME.
DR Genew: HGNC:379; AKAP9.
DR MIM: 604001; -
KW Colled coil; Alternative splicing; Polymorphism.
FT DOMAIN 2554 2567 PKA-RII SUBUNIT BINDING DOMAIN.
FT DOMAIN 164 914
FT DOMAIN 944 1022
FT DOMAIN 1100 1185
FT DOMAIN 1253 1280
FT DOMAIN 1336 1392
FT DOMAIN 1434 1459
FT DOMAIN 1585 1659
FT DOMAIN 1857 2455
FT DOMAIN 2544 2561
FT DOMAIN 2603 2776
FT DOMAIN 3065 3092
FT DOMAIN 3124 3470
FT DOMAIN 3587 3689
FT DOMAIN 3726 3730
FT DOMAIN 203 292
FT DOMAIN 321 1010
FT DOMAIN 1846 2772
FT VARSPLIC 17 28
FT VARSPLIC 1637 1642
FT VARSPLIC 1643 3911
FT VARSPLIC 2175 2182
FT VARSPLIC 2175 2183
FT VARSPLIC 2895 2907
FT VARSPLIC 2895 2948
FT VARSPLIC 3901 3911
FT VARIANT 1347 1347
FT CONFLICT 76 76
FT CONFLICT 475 475
FT CONFLICT 554 554
FT CONFLICT 638 638
FT CONFLICT 663 663
FT CONFLICT 913 913
FT CONFLICT 956 956
FT CONFLICT 980 982
FT CONFLICT 997 997
FT CONFLICT 1001 1001
Q -> Q (IN REF. 3).
M -> I (IN REF. 3).
E -> G (IN REF. 3).
R -> S (IN REF. 3).
N -> S (IN REF. 3).
H -> N (IN REF. 3).
K -> N (IN REF. 3).
OKH -> PKP (IN REF. 1 AND 2).
Q -> P (IN REF. 1 AND 2).
Q -> P (IN REF. 1 AND 2).

FT CONFLICT 1020 1020 N -> D (IN REF. 3).
FT CONFLICT 1028 1028 V -> E (IN REF. 3).
FT CONFLICT 1626 1626 R -> P (IN REF. 1 AND 2).
FT CONFLICT 1703 1703 N -> T (IN REF. 3).
FT CONFLICT 1707 1707 V -> G (IN REF. 3).
FT CONFLICT 1802 1803 MISSING (IN REF. 5).
FT CONFLICT 1843 1843 A -> P (IN REF. 3).
FT CONFLICT 1956 1956 I -> V (IN REF. 3).
FT CONFLICT 2027 2027 V -> D (IN REF. 5).
FT CONFLICT 2157 2158 EI -> HE (IN REF. 7).
FT CONFLICT 2169 2169 E -> V (IN REF. 3).
FT CONFLICT 2514 2514 L -> R (IN REF. 3).
FT CONFLICT 2851 2851 I -> N (IN REF. 8).
FT CONFLICT 2957 2957 E -> D (IN REF. 3).
FT CONFLICT 2983 2983 P -> S (IN REF. 3).
FT CONFLICT 3087 3087 Q -> H (IN REF. 3).
FT CONFLICT 3218 3218 Q -> H (IN REF. 3).
FT CONFLICT 3309 3309 BSE -> OSO (IN REF. 3).
FT CONFLICT 3751 3751 P -> A (IN REF. 3).
FT CONFLICT 3833 3833 T -> S (IN REF. 3).
SQ SEQUENCE 3911 AA; 453664 MM; 3FBICB1C819B47AA CRC64;

Query Match 2.9%; Score 112; DB 1; Length 3911;
Best Local Similarity 20.5%; Pred. No. 28; Mismatches 84; Gaps 15;
Matches 68; Conservative 57; Indels 122; Indels 84; Gaps 15;
QY 11 LDKFHEADLFYQSS---LASWNTNTTEENYNNAGDKWSAFLEKOSTLAQMP 66
DB 1111 LNVLSSEQDRLQWEAQRICLSVYSHVDQR- EYENKDKALGSLKEELFAQEEK 1169
QY 67 LDEIQLN-IVKQLDALQONG-----SSVISEDKSKRLNTILTMSTIYSG 112
DB 1170 IRELQKHOLEQTQTKTQETGEGKPLHLIGLKQKAVSECSYFLQTLCSVLGEYTPA 1229
QY 113 KVC-----NPDRPOCLL-LEPGNEIMANSIDYNE----- 142
DB 1230 LKCEVNAEDKENSQDIYSENEDPELODTRYEVQOEQEMHTLLNKVTEYNNKLVLQTRL 1289
QY 143 -RTAWES--WRSEVGKQLRPLYE-EYVVLKEMARAHYEDYGDYRGDYEVNGVDGYD 198
DB 1290 SKIWGGQDTGKMLERGEENLPKEFEPLSHSQMNL-----DIDVN-----H 1333
QY 199 YRGQLIEVEHT-FEIKPLYLEHAYRAKLMAYSYISPGICLPAHLIGDMGAFW 257
DB 1334 KSKLSLDLEKTKLEQVELESLSISLQOOLKETEONTAEIHCLOKRLQA----- 1386
QY 258 TNLSTVPEFGKPNIDYTDAMVDQAMDAOR 288
DB 1387 --VSESTVP-----PSLPDVSIVITES-DAQR 1410

Search completed: May 26, 2003, 17:54:18
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 16:58:14 ; Search time 69 seconds

(without alignments)
2150.057 Million cell updates/sec

Title: us-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STIEBQAKTFLDKFNHEAD.....DNSLEFLGIQPTLGPNNPP 720

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: sp.archaea:***
2: sp.bacteria:***
3: sp.fungi:***
4: sp.human:***
5: sp.invertebrate:***
6: sp.mammal:***
7: sp.mhc:***
8: sp.organelle:***
9: sp.phage:***
10: sp.plant:***
11: sp.podent:***
12: sp.virus:***
13: sp.vertbrate:***
14: sp.unclassified:***
15: sp.rvivirus:***
16: sp.bacteriaph:***
17: sp.archaeap:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	804	4	Q9UFZ6
2	3869	100.0	805	4	Q9NRA7
3	3866	99.9	805	4	Q9NRA7
4	3251	84.0	798	11	Q9NRA7
5	3251	84.0	805	11	Q9NRA7
6	1461	37.8	353	11	Q9NRA7
7	1329	34.3	732	6	Q9GLN7
8	1329	34.3	1304	6	Q9GLN7
9	1310	33.9	1313	11	Q9EQM9
10	1303	33.7	694	4	Q15540
11	1087	28.1	618	5	Q9NDS8
12	1069	27.6	645	5	Q9NKE4
13	1054	27.2	660	5	Q17248
14	1027.5	26.6	630	5	Q24222
15	1025.5	26.5	630	5	Q9VLJ6
16	930	24.0	844	5	Q8SXX2

17	841	21.7	792	5	Q9VTV1	Q9VTV1 drosophila
18	762	19.7	249	11	Q9DB36	Q9DB36 mus musculus
19	638.5	16.5	907	5	Q18581	Q18581 caenorhabditis
20	491	12.7	661	5	Q9V520	Q9V520 drosophila
21	458	11.8	611	5	Q9VTV2	Q9VTV2 drosophila
22	434.5	11.2	202	11	Q64603	Q64603 rattus norvegicus
23	407.5	10.5	628	5	Q9S079	Q9S079 drosophila
24	407.5	10.5	628	5	Q9W021	Q9W021 drosophila
25	306.5	7.9	121	6	Q9BDG1	Q9BDG1 bos taurus
26	280.5	7.2	135	4	Q16425	Q16425 homo sapiens
27	272	7.0	222	4	Q9HBJ8	Q9HBJ8 homo sapiens
28	271	7.0	222	11	Q9ESG3	Q9ESG3 rattus norvegicus
29	265	6.8	222	11	Q9ESG4	Q9ESG4 mus musculus
30	251.5	6.5	157	11	Q61265	Q61265 mus musculus
31	239	6.2	75	5	Q9TK66	Q9TK66 drosophila
32	157	4.1	532	16	Q9KEVO	Q9KEVO bacillus haemolyticus
33	154	4.0	502	16	Q8Y616	Q8Y616 listeria monocytogenes
34	152.5	3.9	734	5	Q9NKE3	Q9NKE3 drosophila
35	147	3.8	54	6	Q9S161	Q9S161 canis familiaris
36	147	3.8	502	16	Q9ZAC3	Q9ZAC3 listeria monocytogenes
37	139.5	3.6	987	16	Q8TW66	Q8TW66 anabaena sp.
38	139	3.6	608	16	Q9PR80	Q9PR80 ureaplasma urealyticum
39	136	3.5	611	16	Q9PPW8	Q9PPW8 ureaplasma urealyticum
40	130.5	3.4	461	2	Q9K152	Q9K152 enterococcus faecalis
41	125.5	3.2	600	16	Q8R670	Q8R670 fusobacterium nucleatum
42	125	3.2	3571	10	Q9SL27	Q9SL27 arabidopsis thaliana
43	125	3.2	3571	10	Q9SL27	Q9SL27 arabidopsis thaliana
44	124.5	3.2	987	11	Q9AUB4	Q9AUB4 mus musculus
45	124.5	3.2	3429	11	Q08614	Q08614 mus musculus

ALIGNMENTS

RESULT 1

ID	Q9UFZ6	PRELIMINARY;	PRT;	804 AA.
AC	Q9UFZ6			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Hypothetical 92.3 kDa protein (Fragment).			
GN	DKFZ434A014.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TESTIS;			
RA	Wambuit R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;			
RL	Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL110224; CAB53682.1; ..			
DR	MEROPS: M02.006; ..			
DR	InterPro: IPR001548; Peptidase_M2.			
DR	InterPro: IPR000130; Zn_MTPeptidse.			
DR	Pfam: PF01401; Peptidase_M2; 1.			
DR	PRINTS: PR00791; PEPDIPASEA.			
DR	ProDom: PD004184; Peptidase_M2; 1.			
DR	PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.			
KW	Hypothetical protein.			
FT	NON_TER			
SQ	SEQUENCE 804 AA; 92340 MW; 91FF391074CB5DA9 CRC64;			
Query Match	100.0%; Score 3869; DB 4; Length 804;			
Best Local Similarity	100.0%; Pred. No. 6.1e-275;			
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 STIEBQAKTFLDKFNHEADLFYQSLSAAMVNTNTEENVQNNMNGDKWSAFKRGST 60			
DB	18 STIEBQAKTFLDKFNHEADLFYQSLSAAMVNTNTEENVQNNMNGDKWSAFKRGST 77			
QY	61 LAQWPTQIEQNLNVKIQQLALQNGSSVLSSEKSKRLNTLITLMTSTTSGKVCNPDP 120			

```

Db 78 LAQWPIQIOLNLYKQLQALQNGSSVLSSEKSKRLNTILNTSTIYSTGKVCNDNP 137
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGOQLRPLYEEVYLKNEARAHNEH 180
Db 138 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGOQLRPLYEEVYLKNEARAHNEH 197
QY 181 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTEFEIKPLYEHLHAYVAKLMNAPYSISP 240
Db 198 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTEFEIKPLYEHLHAYVAKLMNAPYSISP 257
QY 241 IGLPAHLIGDMGKGFNTLSLYVPGQKPNIDVTAMDVAQADRIFFKAKEFPVS 300
Db 258 IGLPAHLIGDMGKGFNTLSLYVPGQKPNIDVTAMDVAQADRIFFKAKEFPVS 317
QY 301 GLPNMTQGFWENSMITDPGNVOKAVCHPTAMDLSKGFRIILMCTKYTMDFLTAHHEMGH 360
Db 318 GLPNMTQGFWENSMITDPGNVOKAVCHPTAMDLSKGFRIILMCTKYTMDFLTAHHEMGH 377
QY 361 IOYDMAVAAQPLLRNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDFQEDNETINF 420
Db 378 IOYDMAVAAQPLLRNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDFQEDNETINF 437
QY 421 LKQALTYGTLPFTYMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETIC 480
Db 438 LKQALTYGTLPFTYMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETIC 497
QY 481 DPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAAKHEGRLHKCDISNSTEAGOKLFNML 540
Db 498 DPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAAKHEGRLHKCDISNSTEAGOKLFNML 557
QY 541 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLKDQKNKSVGSTMSPYADDSI 600
Db 558 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLKDQKNKSVGSTMSPYADDSI 617
QY 601 KYRISLSKSLGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANLKR 660
Db 618 KYRISLSKSLGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANLKR 677
QY 661 ISFNFTVAPKNVSDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPNO 720
Db 678 ISFNFTVAPKNVSDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPNO 737

RESULT 2
Q9NRA7 PRELIMINARY; PRT; 805 AA.
ID Q9NRA7
AC Q9NRA7:
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 15, Last annotation update)
DE Angiotensin converting enzyme-like protein (ACE2)-related
DE carboxypeptidase ACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RA RA
RC TISSUE=LYMPHOMA;
RA Turner A.J., Hooper N.M., Hyde R.J., Christie G., Kattan E.,
RA "A Human Homolog of Angiotensin Converting Enzyme - Cloning and
RA Functional Expression As A Captopril-Sensitive Carboxypeptidase.";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Donoghue M., Hsieh F., Baronas F., Godbout K., Gosselin M.,
RA Stagliano N., Donovan M., Woolf B., Robison K., Jayaseelan R.,
RA Breitbart R.E., Acton S.;
RA "A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to
RA angiotensin-1-9.";
RL Circ. Res. 0:0-0(2000).

```

```

DR EMBL; AF241254; AAF78220.1; -
DR EMBL; AF291820; AAF99721.1; -
DR MEROPS; M02.006; -
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTPeptidase.
DR Pfam; PF01401; Peptidase_M2; 1.
DR PRINTS; PR00791; PEPDPTASEA.
DR ProDom; PD004184; Peptidase_M2; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Carboxypeptidase.
SQ SEQUENCE 805 AA; 92462 MW; 8EE6EB0A931550B8 CRC64;

Query Match 100.0%; Score 3869; DB 4; Length 805;
Best Local Similarity 100.0%; Pred. No. 6,1e-275;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STIEQATFLDKRNEHEADLEFYOSLSAWNTNTTEENVQNNMNGKSAFLKROST 60
Db 19 STIEQATFLDKRNEHEADLEFYOSLSAWNTNTTEENVQNNMNGKSAFLKROST 78
QY 61 LAQWPIQIOLNLYKQLQALQNGSSVLSSEKSKRLNTILNTSTIYSTGKVCNDNP 120
Db 79 LAQWPIQIOLNLYKQLQALQNGSSVLSSEKSKRLNTILNTSTIYSTGKVCNDNP 138
QY 121 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGOQLRPLYEEVYLKNEARAHNEH 180
Db 139 QECILLEPGLNEIMANSIDYNERLWAMESRSEVGOQLRPLYEEVYLKNEARAHNEH 198
QY 181 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTEFEIKPLYEHLHAYVAKLMNAPYSISP 240
Db 199 YGDYWRGDEYVNGVGYDYSRQGLIEDVEHTEFEIKPLYEHLHAYVAKLMNAPYSISP 258
QY 241 IGLPAHLIGDMGKGFNTLSLYVPGQKPNIDVTAMDVAQADRIFFKAKEFPVS 300
Db 259 IGLPAHLIGDMGKGFNTLSLYVPGQKPNIDVTAMDVAQADRIFFKAKEFPVS 318
QY 301 GLPNMTQGFWENSMITDPGNVOKAVCHPTAMDLSKGFRIILMCTKYTMDFLTAHHEMGH 360
Db 319 GLPNMTQGFWENSMITDPGNVOKAVCHPTAMDLSKGFRIILMCTKYTMDFLTAHHEMGH 378
QY 361 IOYDMAVAAQPLLRNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDFQEDNETINF 420
Db 379 IOYDMAVAAQPLLRNGANGEGHEAVGEIMSLSAATPKHLKSGILSPDFQEDNETINF 438
QY 421 LKQALTYGTLPFTYMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETIC 480
Db 439 LKQALTYGTLPFTYMLEKRMWVKGELPKDQMKKMKMKREIYGVVEVPHDETIC 498
QY 481 DPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAAKHEGRLHKCDISNSTEAGOKLFNML 540
Db 499 DPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAAKHEGRLHKCDISNSTEAGOKLFNML 558
QY 541 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLKDQKNKSVGSTMSPYADDSI 600
Db 559 RLGKSEPTLALENVGAKNMNVRPLNTEPELFTWLKDQKNKSVGSTMSPYADDSI 618
QY 601 KYRISLSKSLGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANLKR 660
Db 619 KYRISLSKSLGDKAYEMNDENKYLFRSSVAYAMROYFLKYNOMILFGEEDVYVANLKR 678
QY 661 ISFNFTVAPKNVSDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPNO 720
Db 679 ISFNFTVAPKNVSDIIPTEVEKAIKMSRSRINDAFRLNDSLEFLGIQPLGPNO 738

RESULT 3
Q9BYF1 PRELIMINARY; PRT; 805 AA.
ID Q9BYF1
AC Q9BYF1:
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ACE2.

```

GN ACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Komatsu T., Suzuki Y., Sugano S.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Suzuki Y., Watanabe M., Sugano S.;
RT Cloning, expression analysis and chromosomal localization of a novel
RT ACE like enzyme."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046569; BAB40370.1; .
DR MEROPS: M02.006; .
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR001530; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPTIDASE.
DR ProDom: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 805 AA; 92491 MW; D2AAB4C27088B72 CRC64;

Query Match 99.9%; Score 3866; DB 4; Length 805;
Best Local Similarity 99.9%; Pred. No. 1e-274;
Matches 719; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 STIEOAKTEFLDKFHEAEEDLFYSSLSASWNTNTTEENVONMNNAGDKMSAFLEKQST 60
19 STIEOAKTEFLDKFHEAEEDLFYSSLSASWNTNTTEENVONMNNAGDKMSAFLEKQST 78
61 LAOWYPLQEIQLNLTIVKIQALQALQNGSSVLSSEDSKRLNTILNTMTSTIYSGKVCNPDP 120
79 LAOWYPLQEIQLNLTIVKIQALQALQNGSSVLSSEDSKRLNTILNTMTSTIYSGKVCNPDP 138
121 QECILLEPGLNEIMANSLDYNERLMAWESRSEVGKOLRPLYEEVYLKEMARAHYED 180
139 QECILLEPGLNEIMANSLDYNERLMAWESRSEVGKOLRPLYEEVYLKEMARAHYED 198
181 YGDYWRGDEYNGVDGYDSRGQLIEDVEHTEFEETKPLYEHLHAYRAKLANAVPSYIS 240
199 YGDYWRGDEYNGVDGYDSRGQLIEDVEHTEFEETKPLYEHLHAYRAKLANAVPSYIS 258
241 IGCPLAHLGDMWGFRFTNLISLTVPFGOKPNIDVTAMVQADAMQRIKFAEKFEVSV 318
259 IGCPLAHLGDMWGFRFTNLISLTVPFGOKPNIDVTAMVQADAMQRIKFAEKFEVSV 336
301 GLPNTQGFWEENSMITDPGNOKAVCHPTAMDLSGDFRIILCTVYMDDEFLTAHENG 360
319 GLPNTQGFWEENSMITDPGNOKAVCHPTAMDLSGDFRIILCTVYMDDEFLTAHENG 378
361 IGYDMAVAAOPPLRLNGANGEGFHEAVGELTMSLSAATPKHLKISIGLSDFQEDNETEIN 420
379 IGYDMAVAAOPPLRLNGANGEGFHEAVGELTMSLSAATPKHLKISIGLSDFQEDNETEIN 438
421 LKQALITVGLPTFTYMLEKRWAVFKGEIPKIDQMKMKWEKREIYGVVPRPHDETTC 480
439 LKQALITVGLPTFTYMLEKRWAVFKGEIPKIDQMKMKWEKREIYGVVPRPHDETTC 498
481 DPASLFEHNSNDYSFTRYYTRRLYQPOFQALCOAAKHGPHKCDINSSTAGOKLNM 540
499 DPASLFEHNSNDYSFTRYYTRRLYQPOFQALCOAAKHGPHKCDINSSTAGOKLNM 558
541 RLKSEPTLALENVYGAKKNNVRLNLTFFLTWTLDQKNKSFVQSTWSPYADQSI 600
559 RLKSEPTLALENVYGAKKNNVRLNLTFFLTWTLDQKNKSFVQSTWSPYADQSI 618
601 KYRISLKSALDQKAVEMNDNEMYLEFRSSVAYAMROYFLAYKNOMILLEGEDVYANIKPR 660
619 KYRISLKSALDQKAVEMNDNEMYLEFRSSVAYAMROYFLAYKNOMILLEGEDVYANIKPR 678

661 ISNFEVYAPKNVSDIIPTEVEKATIRMSRINDAFRLNDSLEFLGIQPTLGPDPNP 720
679 ISNFEVYAPKNVSDIIPTEVEKATIRMSRINDAFRLNDSLEFLGIQPTLGPDPNP 738

RESULT 4
ID 099N71 PRELIMINARY; PRT; 798 AA.
AC 099N71;
DT 01-JUN-2001 (TRENBERG, 17, Created)
DT 01-JUN-2001 (TRENBERG, 17, Last sequence update)
DT 01-JUN-2002 (TRENBERG, 21, Last annotation update)
DE Angiotensin-converting enzyme-related carboxypeptidase.
GN 2010305105RIK OR ACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Ymada Y.,
RA Hida M., Tanigami A., Muroi S.;
RT Molecular cloning, mRNA expression, and chromosomal localization of
RT mouse Angiotensin-converting enzyme-related Carboxypeptidase."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB053181; BAB40371.1; .
DR MEROPS: M02.006; .
DR MGD: MG1:1917258; 2010305105RIK.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR001680; W040.
DR InterPro: IPR001530; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPTIDASE.
DR ProDom: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Carboxypeptidase.
SQ SEQUENCE 798 AA; 91943 MW; 403AEA2D55725A4 CRC64;

Query Match 84.0%; Score 3251; DB 11; Length 798;
Best Local Similarity 83.18%; Pred. No. 1.2e-229;
Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

1 STIEOAKTEFLDKFHEAEEDLFYSSLSASWNTNTTEENVONMNNAGDKMSAFLEKQST 60
19 STIEOAKTEFLDKFHEAEEDLFYSSLSASWNTNTTEENVONMNNAGDKMSAFLEKQST 78
61 LAOWYPLQEIQLNLTIVKIQALQALQNGSSVLSSEDSKRLNTILNTMTSTIYSGKVCNPDP 120
79 LAOWYPLQEIQLNLTIVKIQALQALQNGSSVLSSEDSKRLNTILNTMTSTIYSGKVCNPDP 138
121 QECILLEPGLNEIMANSLDYNERLMAWESRSEVGKOLRPLYEEVYLKEMARAHYED 180
139 QECILLEPGLNEIMANSLDYNERLMAWESRSEVGKOLRPLYEEVYLKEMARAHYED 198
181 YGDYWRGDEYNGVDGYDSRGQLIEDVEHTEFEETKPLYEHLHAYRAKLANAVPSYIS 240
199 YGDYWRGDEYNGVDGYDSRGQLIEDVEHTEFEETKPLYEHLHAYRAKLANAVPSYIS 258
241 IGCPLAHLGDMWGFRFTNLISLTVPFGOKPNIDVTAMVQADAMQRIKFAEKFEVSV 318
259 IGCPLAHLGDMWGFRFTNLISLTVPFGOKPNIDVTAMVQADAMQRIKFAEKFEVSV 336
301 GLPNTQGFWEENSMITDPGNOKAVCHPTAMDLSGDFRIILCTVYMDDEFLTAHENG 360
319 GLPNTQGFWEENSMITDPGNOKAVCHPTAMDLSGDFRIILCTVYMDDEFLTAHENG 378
361 IGYDMAVAAOPPLRLNGANGEGFHEAVGELTMSLSAATPKHLKISIGLSDFQEDNETEIN 420
379 IGYDMAVAAOPPLRLNGANGEGFHEAVGELTMSLSAATPKHLKISIGLSDFQEDNETEIN 438
421 LKQALITVGLPTFTYMLEKRWAVFKGEIPKIDQMKMKWEKREIYGVVPRPHDETTC 480

Db 439 LKQALLTVGLPTFTYMLEKRWMMVFGEIPKEQMMKKWMEKREIVGVPELPHDITYC 498
 QY 481 DPASLEFHSNDYSFTRYYTRRLYOFOFOBALCOAAKHEGPIHKDCDISNSTEAGOKLENNM 540
 Db 499 DPASLEFHSNDYSFTRYYTRRLYOFOFOBALCOAAKHEGPIHKDCDISNSTEAGOKLENNM 558
 QY 541 RLKSESPWTLALENVGAKNNMVRPLNYFEPLFTWLKQKNSFVGNSTDSWSPYADQSI 600
 Db 559 SIGNSEPWTKALENVGAKNNMVRPLNYFEPLFTWLKQKNSFVGNSTDSWSPYADQSI 618
 QY 601 KVRISLKSALGDKAYEWMNDENMYLFRSSVAYAMROYFLKYNOMLFGEDVRYANLKPR 660
 Db 619 KVRISLKSALGANAAYEWNNEMFLFRSSVAYAMRYEFTIKNQVPLEEDVRYSDLKPR 678
 QY 661 ISPNFVTPAPKNVSDIIPRETEVEKARMSRSRINDAFRLNDSLEFLGIQPTLGPNNP 720
 Db 679 VSEFFVTSPQNVSDVYIPRSEVEDALRMSRGRINDVFGINDNSLEFLGIHPTLEPPYOP 738

RESULT 5

Q8R010 PRELIMINARY; PRT; 805 AA.
 ID Q8R010: 01-JUN-2002 (TREMBlrel. 21, Created)
 AC 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE RIKEN CDNA 2010305L05 gene.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strusberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC026801; AAH26801.1;
 SQ SEQUENCE 805 AA; 92367 MW; DB883AAC966A8D9 CRC64;

Query Match

Best Local Similarity 83.1%; Pred. No. 1.2e-229; Length 805;
 Matches 598; Conservative 49; Mismatches 73; Indels 0; Gaps 0;

1 STIEBOAKFLDKFHEAEDLFYQSSLASWNTNITEENYQNNNGDKWSAFLEKQST 60
 Db 19 SLTEENAKFLNNFNOEAEDLSYQSSLASWNTNITEENYQNNNGDKWSAFLEKQST 78
 QY 61 LAOMTPELONTLVKLOALQOQSSSVLSSEDSKRLNTLNTSTYSGKVCNPNP 120
 Db 79 TAQSFSLQETQPTIKROLQALQOQSSSALSDKKOKLNTLNTSTYSGKVCNPNP 138
 QY 121 QECILLEPGINEIMANSIDYNERLWAMESWSEVQKOLPLYLEEYVVLKNNMAYRHHED 180
 Db 139 QECILLEPGIDEIMASTDYNSRLWAMEGMALEVQKOLPLYLEEYVVLKNNMAYRHHED 198
 QY 181 YGDYWRGDYEVNGVDGYDYSRGQILIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 240
 Db 199 YGDYWRGDYEAEGADGYNNRNQOLIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 258
 QY 241 IGLPLPHLLGDWNGREFTNLYSLTFVFGQKPNIDYTDAMVDQAMDQRIFEAKKEFYVS 300
 Db 259 TGCPLPHLLGDWNGREFTNLYSLTFVFGQKPNIDYTDAMVDQAMDQRIFEAKKEFYVS 318
 QY 301 GLPMMTQGFWNSMLTDPGNVOKAVCHPAPMDLQKDRILMCTVYTMDDPLTAHHEKGH 360
 Db 319 GLPMMTQGFWNSMLTDPGNVOKAVCHPAPMDLQKDRILMCTVYTMDDPLTAHHEKGH 378
 QY 361 IQYMAAYAPPELIRNANEGFHEAYGEIEMSLAATPKHLKSLIGLLSPFOEDNETELNF 420
 Db 379 IQYMAAYAPPELIRNANEGFHEAYGEIEMSLAATPKHLKSLIGLLSPFOEDNETELNF 438
 QY 421 LKQALLTVGLPTFTYMLEKRWMMVFGEIPKEQMMKKWMEKREIVGVPELPHDITYC 480

RESULT 6

Q99N70 PRELIMINARY; PRT; 353 AA.
 ID Q99N70: 01-JUN-2001 (TREMBlrel. 17, Created)
 AC Q99N70: 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Angiotensin-converting enzyme-related carboxypeptidase.
 GN ACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Yamada Y.,
 RA Hide M., Tanigami A., Muroi S.;
 RT "Molecular cloning, mRNA expression, and chromosomal localization of
 RT mouse Angiotensin-converting Enzyme-Related Carboxypeptidase."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

121 QECILLEPGINEIMANSIDYNERLWAMESWSEVQKOLPLYLEEYVVLKNNMAYRHHED 180
 Db 139 QECILLEPGIDEIMASTDYNSRLWAMEGMALEVQKOLPLYLEEYVVLKNNMAYRHHED 198
 QY 181 YGDYWRGDYEVNGVDGYDYSRGQILIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 240
 Db 199 YGDYWRGDYEAEGADGYNNRNQOLIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 258
 QY 241 IGLPLPHLLGDWNGREFTNLYSLTFVFGQKPNIDYTDAMVDQAMDQRIFEAKKEFYVS 300
 Db 259 TGCPLPHLLGDWNGREFTNLYSLTFVFGQKPNIDYTDAMVDQAMDQRIFEAKKEFYVS 318
 QY 301 GLPMMTQGFWNSMLTDPGNVOKAVCHPAPMDLQKDRILMCTVYTMDDPLTAHHEKGH 360
 Db 319 GLPMMTQGFWNSMLTDPGNVOKAVCHPAPMDLQKDRILMCTVYTMDDPLTAHHEKGH 378
 QY 361 IQYMAAYAPPELIRNANEGFHEAYGEIEMSLAATPKHLKSLIGLLSPFOEDNETELNF 420
 Db 379 IQYMAAYAPPELIRNANEGFHEAYGEIEMSLAATPKHLKSLIGLLSPFOEDNETELNF 438
 QY 421 LKQALLTVGLPTFTYMLEKRWMMVFGEIPKEQMMKKWMEKREIVGVPELPHDITYC 480

Query Match

Best Local Similarity 37.8%; Score 1461; DB 11; Length 353;
 Matches 268; Conservative 22; Mismatches 43; Indels 0; Gaps 0;

1 STIEBOAKFLDKFHEAEDLFYQSSLASWNTNITEENYQNNNGDKWSAFLEKQST 60
 Db 19 SLTEENAKFLNNFNOEAEDLSYQSSLASWNTNITEENYQNNNGDKWSAFLEKQST 78
 QY 61 LAOMTPELONTLVKLOALQOQSSSVLSSEDSKRLNTLNTSTYSGKVCNPNP 120
 Db 79 TAQSFSLQETQPTIKROLQALQOQSSSALSDKKOKLNTLNTSTYSGKVCNPNP 138
 QY 121 QECILLEPGINEIMANSIDYNERLWAMESWSEVQKOLPLYLEEYVVLKNNMAYRHHED 180
 Db 139 QECILLEPGIDEIMASTDYNSRLWAMEGMALEVQKOLPLYLEEYVVLKNNMAYRHHED 198
 QY 181 YGDYWRGDYEVNGVDGYDYSRGQILIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 240
 Db 199 YGDYWRGDYEAEGADGYNNRNQOLIEDVHTEFEIKPLYEHLHAYVRKAKNNAPSYISP 258

Query Match	33.7%;	Score 1303;	DB 4;	Length 694;
Best Local Similarity	42.7%;	Pred. No. 5.7e-87;		
Matches 250;	Conservative 107;	Mismatches 105;		

```

QY 27 LASNYNTNITTEE-----NYONNNAAGDKWASFLKJQSTLAQMYPLQEOJNTLTKRQ 78
Db 57 IATNYNTNITTEFSKILLQKNQJANH-----LKQJQARKEFDVNOJNTTKRI 108
QY 79 LQALQNGSSVLSSEDSKRLNTLNTNMSTYTSQKVCNDNQECLLEPGLNEMANSL 138
Db 109 IKKQODLERAAALPAQOELEENKILLDMETTSYATVCHPNQ-----SCJQLEPDLTNMASTR 166
QY 139 DYNBRMAMWSMSEVQKOLRPLYEETVYVKNEMARANHEDYGDYWRGDEYVANGVDGD 198
Db 167 KYEDLMAWGMKDKKRAILLQFPYKVELINQARLNQYDAGSWMSMETPBLE---- 223
QY 199 YSRGOLIEDVEHTEFERIKPLYEHLTAHYVAKIMNAY-PSYSPISICLPAHLGDMGKRFW 257
Db 224 -----QDLERFQELQPLYNLHAYRRAALHRRYGAOHNLBESPIHAHLGNMAMQTW 276
QY 258 TNLSTLYVEFGQKPNIDYTDANVDQAMDQRIETKEAEKFVYSGVJPNNTQGFWMSMTJD 317
Db 277 SNITDLYVPEPSAPSDTTEAMLKQGMWPRRFEKADDFETSLGILPVPPEFKMSLKE 336
QY 318 PSNVQAVACHPTAMDLGKG-DFRIMACTKRYMDEDTLHNHMGHQQDMAVYAAOPFLRN 376
Db 337 PJDGGEVYVCHASAMDFNGKDFRIKQCTVLMEDLYVAHHBMGHQYFMQYKDPJVALRE 396
QY 377 GANEGFHVAQVIMLSATPRKHLKSGILSPDFQEDNETEINTLQALTYIVGTLPETY 436
Db 397 GANPEFHHAIGVALLSVSTPRKHLHSILNISEGSGD-EHDINFLMMLDKIAFIPEST 455
QY 437 MLEKRWAVYFGKEJPKQDMQMKWMEKREIYGVYVPEPHDQTYDDPALSLEHVSNDSFIR 496
Db 456 LVDDQMRVHVFQDGSITKENYQEMWSLKLKQGLCPVPRQOGDPPAKHIPPSSVPIYR 515
QY 497 YUTRILYFOFOEALCOAKHEGPLHKODISNSTEAGOKLFNMLRLKSESPFATLENY 556
Db 516 YFVSFTILOFOHEALCOAAGHTGFLHKCDIYQSKRGAORLATAMKIGFSPRWEANOLIT 575
QY 557 GAKNMNVRPLNTEPEPLFTMLQDNK-----NSFYGM-SIDWSPRYADQS 599
Db 576 GQPMNASASAMLSYFKRPLDMDKRTENELHGEKTLQMPQOINTMPPNARS 621

```

RESULT	11		
Q9NDS8	Q9NDS8	PRELIMINARY;	PRT; 648 AA.
ID	Q9NDS8		
AC	Q9NDS8		
DT	01-OCT-2000	(REMBREL. 15	Created)
DT	01-OCT-2000	(REMBREL. 15	Last sequence update)
DT	01-DEC-2001	(REMBREL. 19,	Last annotation update)

Query Match	28.1%	Score 1087	DB 5	Length 648
Best Local Similarity	36.1%	Pred. No. 3.5e-71		
Matches 228	Conservative 122	Mismatches 248	Indels 34	Gaps 10

```

0Y 1L LDRKNHAEDELFT-----OSSLASWNTNITTEENVONMMNADCKWASFLKQGS 59
Db 29 LEAREHAREYMLHLDRKATGRLKRRNRASLAEWETYSNITKENBEKSIQTLIELTSROEKAW 88
0Y 60 TLAMPTLOEIONLTVKLOLQALQNGSSVLSDESKRLNTILNMTSTISGKACPDN 119
Db 89 EEIKRMGWOPFOFTLRMRKYSQIGVALALDDXFOALMRVSOESNVAFAKCSYKN 148
0Y 120 POEC-LTLEBGLMEIANSIDYERLIMANESRSEYGOALRELYEEVVLAKENARAHY 178
Db 149 ESKCDLSLEPEITELTISTODPEELKHAWEHNHNAAGATANKFIDYVNLNEAKLNGE 208
0Y 179 EDYGDWRGDEYEVNGDYDYSRGOLIEDVEHTEFEIKPLIEHLYAVRAKLYMAY-PSY 237
Db 209 DVAEWMOSEYEVPDE-----EOLAKTMEVAPLVOQLHAYRKRLRKTYDXY 258
0Y 238 ISEIGCLPAHLIGDMXGFWNTLYSLTVPQGRKPIDVTDYAMQANDAOIRFEAREKF 297
Db 259 VSARGPILPAHLGNNMAOTWNNIEFTPRYPDKRKTDYQAMROUNTPMKMEPMSDEFF 318
0Y 298 VASGJPLNMGOEWMSMLTDGQNVOKAVCHPTANDLGK-DFTIMCTKYMTDFTLAH 356
Db 319 RSLNLTAMPEKWKMSIIEKPTD-REIYCHASANDFEGGEFPERIKOCTVDEYEQTHH 377
0Y 357 EKHGIQYDMAVAADPELLRNGANBGFHVAEIGMSLSAUPRKHLSIGLSPDOENEG 416
Db 378 EMGHIOYLYQREDQVYPERDQANGFHEVADTIALSVSSPKHLRRGLATYG-ADDEOT 436
0Y 417 EINFILKQALTVGLPFTYMLEKMRMAYFRGELPRKQOMKKMKMKREIYGVVEVPDH 476
Db 437 EITNQYKNGIDKIAFLPPAYITLDFRGVFRFRKRLPIDYCHWKLRLOJQGEPPVNR 496
0Y 477 EYCDPASLFWVSDYSFIRYRTTYLQFOFQALCO-AAKHEG-----PLHKCDISNST 530
Db 497 EDDPDAKAHYVSSNVEYARYVSFIIOFOFHGVCQLAGEHAADGPNKKIKVDCOLYGSV 556
0Y 551 EAGKLEFNLRLGSEPEWTLAEVYVAKAMNVRPLINTPEELFTYLKQONKS--FVGM 588
Db 557 AAGNALANMLKMGSSKPPMDAMEALTOGRKMKADGILETFRPLHDLTRAENORTGEHIGW 616
0Y 589 STDWSPADQSIKTRISLKSALDGKAWENDN 620
Db 617 EPTMEECTPSQSLSELNWKPESSSPATQSDS 648

```

RESULT 12

09NKE4 PRELIMINARY: PRF: 615 AA.

ID 09NKE4; 09YTV3; Created)

DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DR Ance protein (LJ1258P).

GN ANCE OR CG8827.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE=99403001; PubMed=10471707;

RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazek R., Davis T., Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D., Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A., Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K., Celinker S., Rubin G.M.;

RA "An exploration of the sequence of a 2.9-Mb region of the genome of Drosophila melanogaster: the Adh region.";

RT Genetics 153:179-219(1999).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RX MEDLINE=201996006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H.C., Blazek R.G., Champagne M., Pfeiffer B.D., Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bertone P., Bickel J., Brokstein P., Brotler P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Chervak D., Chokyan M.R., Bouck J., Brockstein P., Brothier P., Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de la Cruz A., Deng Z., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson B., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W., Fodor C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K., Glisler A., Gong F., Gottlieb J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C., Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Lammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulyov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhong L., Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RA "The genome sequence of Drosophila melanogaster.";

RT Science 287:2185-2195(2000).

RL [4]

RN SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J., Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nuno J., Pacle J., Parasas V., Park S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celinker S.;

RA Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.

RL EMBL: AE003408; AAF44834.1; -

DR EMBL: AE003641; AAF53353.2; -

DR EMBL: AY061129; AAL28677.1; -

DR FlyBase: FBgn0012037; Ance.

DR InterPro: IPR001548; Peptidease_M2.

DR InterPro: IPR001130; Zn_MTPeptidease.

DR Pfam: PF01401; Peptidease_M2; 1.

DR PRINTS: PR00791; PEPTIDASEA.

DR ProDom: PD004184; Peptidease_M2; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.

DR SEQUENCE 615 AA; 70913 MW; 9E3691BC51D6C48 CRC64;

Query Match 27.64; Score 1069; DB 5; Length 615;

Best Local Similarity 36.38; Pred. No. 6,8e-70;

Matches 217; Conservative 120; Mismatches 239; Indels 22; Gaps 9;

4 EQAKTLEDFKNEAEDELFYQSSIASWNTNTTEENQONMNAAGKWSAFKKEOSTLAQ 63

22 EIQKKEYLENKLKELARTVETEAAYAGSNTDENKKKISAEIAKFEVASTDT 81

64 MYPLQELONLTVKLOIALOONGSSVSEDSKRLMTILNTSTYSTGVCNPNPQRC 123

82 KFWKSTYQSEDLKQFALTKLGYALPBDYAEILDITLSAMSPNAKVCYCDKSTKC 141

124 -LLEPGELNEIANSIDYNERLWAMESRSEVQKOLPLVEEYVYKNEARHNHEDG 182

142 DLALDPEIEEYVSKSRDHEELAYWRBFFYKAGRAVVSQERVEYVNTKAKLNTTSSA 201

183 DYWRDYEYVNGVDYDYSRGOLIDVETFEIEIKPLEYHAAVYRAKLNAV-PSYISPI 241

202 EAWIDEYE-----DITFEQLEDI---FADIRPLVQIGHYRFRKHYGDVAVSET 251

242 GCLPAHLIGDMWRGFWTNIYSILVPGGKPNIDVTAMDADQADQKIFKEAKFFVSVG 301

252 GFLPHLLGNMMAQOMSEIADIVSPFPEKPLVDVSAEKEKGYPPLMFQMGDPFTSMN 311

302 LPMATGQFENEMLDLPGVQVQKACHPTAMDG-KGDFRLIMCTKVTMDFLAHHEMGH 360

312 LKRLPDQFEDKSLITERPDGRLVCHASAMDFYLDVIRICGRVQDQLFVHRELGH 371

361 IQYDMAYAAPPLLRNGANGFHEAVGEIMTSIAATPKHLKSLGSLSPDQEDNETINF 420

372 IQYFLQYQHPFYRRGANGFHEAVGDLSTLSTVSPKHEKIGLKL-DYVRDDEARINQ 430

421 LKQALITVGLTFLPTMLKRWMPYKGIPIPDQMKKRWMEKKRIYGVVPHDETTC 480

431 LFTALDKVFLFPAFTMKYRNSLFRGSDVDRANKNCAFWKLRDYSIEPPVRSKEDF 490

481 DPASLTVSNDYSFIRYTRTYLQFOFQALC-QAKH-----EGPLHKDISNTEGQ 534

491 DAPAKYIISDVETKIVISFIQFQYKACIKAGQYDPDVAEPLDNCIDYGSAAAGA 550

535 KLFNMLRLGSEPTLALENVVAKAMNVRPLNTPEPLFTWLKQON--KNSFVMSR 590

551 AFHMLSMGASKIPDPALEAFNGERIMSGALAEYFPLRWLEAKNKNNYHIGMT 608

QY 304 NMTGCFWNSMLTDGNOVAKVCHPTANDLCK-GDFRIMCTKYTMDFLHHEMHQ 362
 DB 323 ALPSEFWMLSVLTRDD--RQVVCASANDFYQDSVVRKMKCTEVDSHFYVHHHLHIQ 381
 QY 363 YMAAFAOPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGSLSPDFQEDNETEINFL 422
 DB 382 YLQYEQDPAVYRGAPNPGFHEAVGDVIALSVMSKHLKAGLIE-NORLDEKSRINOLF 440
 QY 423 KQALITVGLPTPTYLEKRWNVFEGELPKDQMKMKEMKREIVGVYVPPHDEYCDP 482
 DB 441 KQALSKIVLPFGVADKYRAVFNELDESQMGNGFQOMSEFGVYVPPHDEYCDP 500
 QY 483 ASLFFVNSDYSTRYRTYTRTLVQFOFQALCOAKHEGP-----LHKCDISNSTEAGOKL 536
 DB 501 PAKYHIDADVETLRFAHIFOFHFKALCRAGOTAPRNSLTLTDNCIDIGSKRAAGSL 560
 QY 537 FNNLRIGKSEPTTALLENVVGAKNNVREPLNYFEPLTWLKDQKNSFVGSTWSP 594
 DB 561 SQFLSGNSRHMKEVLEETGETEMDPALLLEYFPLVQWLKOE--NSRLGVLGMP 616

RESULT 15
 QYVLJ6 PRELIMINARY; PRT; 630 AA.
 ID 09VLJ6
 AC 09VLJ6;
 DT 01-MAY-2000 (TREMELREL. 13, Created)
 DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)
 DT 01-DEC-2001 (TREMELREL. 19, Last annotation update)
 DE ACER protein (J028328P).
 GN ACER OR CG10593.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Phnydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos B.L.,
 RA April J.F., Abghyani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bayandale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Gelbart W.M., Glasser K.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelpi D., Gish W.B., Glaser K.,
 RA Glodde K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Heston D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laako P., Lei Y., Levitsky A.C., Li J., Li Z., Lian X.,
 RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shie B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhibo X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RN Science 287:2185-2195(2000).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Partan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Paele J., Parag V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003621; AAF52693.1;
 DR EMBL: AY051750; AAK93174.1;
 DR MEROPS: M02.002;
 DR FLYBASE: FBgn0016122; Acer.
 DR InterPro: IPR001548; Peptidase_M2.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPDIPASEA.
 DR ProDom: PD004184; Peptidase_M2; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN 1.
 SQ. SEQUENCE 630 AA; 73057 MW; 6D9355EB5773289 CRC64;

Query Match 26.5%; Score 1025.5; DB 5; Length 630;
 Best Local Similarity 36.0%; Pred. No. 1.1e-66;
 Matches 215; Conservative 109; Mismatches 251; Indels 23; Gaps 9;

QY 6 QAKTFLEKFNEDDLDFYQSSLASMYNTNTTEENYONMNNAGDKSAFIKEGSTLAOMY 65
 DB 33 EARRFELENNQLRRRHEEFLSGYNTNTTEENYONMNNAGDKSAFIKEGSTLAOMY 92
 QY 66 PLOEIONLTVKIQLOALQONGSSVLSSEKSKRLTILNTSTIYSGKVGPNPQPC-L 124
 DB 93 DYQSESDADIRQAEHLKSLGASALNMDVIALQNALISSQGTNAATVCSYTRBSCSL 152
 QY 125 LLEPGLNFIAMSLDYNERNLWMSWSENGKOLRPLERYVYLKMKMARNEYDGY 184
 DB 153 TLEPHIQRLSHSRPALAWYRWEMHDKSTPKRONAEVRLTRRAASLNGHRSYADY 212
 QY 185 WRGDEVGVGVGYDVSQGLLEDEYHPEETKPLYEHLHAYRAKLNNAY-PSYISPIGC 243
 DB 213 WQVTFE-----DPPFER-----QIDATPKQLPFRQLHGVRRRLAQHGPDVMPAEGN 262
 QY 244 LPAHLIGDMGRFTWNLISLTVPGQKPNIDVTDAVDAQDAORIFKEAKEFVSGLP 303
 DB 263 IPISLGNMGWGSNNELDLFTPYPEKPFVDVKAEMEKGYVOKLELGDQFOSLGM 322
 QY 304 NMTGCFWNSMLTDGNOVAKVCHPTANDLCK-GDFRIMCTKYTMDFLHHEMHQ 362
 DB 323 ALPSEFWMLSVLTRDD--RQVVCASANDFYQDSVVRKMKCTEVDSHFYVHHHLHIQ 381
 QY 363 YMAAFAOPFLIRNGANGFHEAVGEIMSLSAATPKHLKSLGSLSPDFQEDNETEINFL 422
 DB 382 YLQYEQDPAVYRGAPNPGFHEAVGDVIALSVMSKHLKAGLIE-NORLDEKSRINOLF 440
 QY 423 KQALITVGLPTPTYLEKRWNVFEGELPKDQMKMKEMKREIVGVYVPPHDEYCDP 482
 DB 441 KQALSKIVLPFGVADKYRAVFNELDESQMGNGFQOMSEFGVYVPPHDEYCDP 500
 QY 483 ASLFFVNSDYSTRYRTYTRTLVQFOFQALCOAKHEGP-----LHKCDISNSTEAGOKL 536
 DB 501 PAKYHIDADVETLRFAHIFOFHFKALCRAGOTAPRNSLTLTDNCIDIGSKRAAGSL 560
 QY 537 FNNLRIGKSEPTTALLENVVGAKNNVREPLNYFEPLTWLKDQKNSFVGSTWSP 594
 DB 561 SQFLSGNSRHMKEVLEETGETEMDPALLLEYFPLVQWLKOE--NSRLGVLGMP 616

RESULT 16
 Q8SXX2

ID 085XX2 PRELIMINARY; PRT; 844 AA.
 AC 085XX2: PRELIMINARY; PRT; 844 AA.
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RH06639P.
 GN ANCE-3
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phuanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Cejner S.,
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AY075532; AL68339.1;
 DR SEQUENCE 844 AA; 98243 MW; 46E1E123F5C8F71 CRC64;

Query Match 24.0%; Score 930; DB 5; Length 844;
 Best Local Similarity 33.2%; Pred. No. 1.7e-59;
 Matches 194; Conservative 119; Mismatches 248; Indels 24; Gaps 9;

QY 28 ASNNYNTITEENVQNNMNNAGDKMSAFLEQSTLAOMYPLDIOQLYKLOLQALQNGS 87
 DB 234 AQNNFETNVNDFQTEALNANOQRYVEFGQITADQSRINKDLIFRRLYROLMOSEVG 293
 QY 88 SYLSEKSRKNTIITASTIYSTGVCPNDPQC-LILEPGLNEIMANSIDYNERLMA 146
 DB 294 NALPLDYLDKYNLEMLFLYNSAFICAYQOPFCODLHYIQLKDIAKSDMELOHT 353
 QY 147 MESWSEVQKOLRPLTEEVVYKNEKARANEHEDGDRWGDYENGVGDYDSGQILE 206
 DB 354 WVEYHKKAGRGKNDSEYGLIDWQEVAVYNNNTNGERTYLAIE-----SGNFRQ 403
 QY 207 DVNHTFEETKPLYLEHATVYAKLMANAY-PSYISPGCLPAHLGDMGRFWTNYSLTY 265
 DB 404 DMQIVNQIRPLIEGLHAYVRKRLDYGPDRINRIAPISPHILGNMYGQSMNVLDILI 463
 QY 266 PEGQKPNIDVAMVDAQDAORIRKEKEKFEVSGLEKMTQGFRENSMLTDCGVQKAV 325
 DB 464 PYRGRLIDVPRMVEQGTTPOLMQLAEFEFTSNMSAVGPFEYRNSIFRQPID-RVLL 522
 QY 326 CHPTAMD-L-GKGFRLIMCTKVTMDFLTAHHEMGIQDYMAVAAOPLLRNGNEGHE 384
 DB 523 CEPASAMDFCNRRHDFVTKICTDINQSLISVHEMAHIOYFYQYHILKRIFRNGANPFRHQ 582
 QY 385 AVGEIMLSAATPKRLKSGISLSPFQDNTEINFLIKQALTIVGLPFTYMLEKRW 444
 DB 583 AVGDIAIGSVSTPRKLTGLQGRSLDE-SSYDINLYETMAIDVAVLPELALSDNRWD 641
 QY 445 VKGEIPIKDWKMKWMEKREIVGVVEPYDEYTCOPASIFHSNDYSFRYTRTLXQ 504
 DB 642 VFSGANRRTMNCCHVMNRKESYSGKPYVLASEDFDCAHYHPIAPIPYIKYFFSTVQ 701
 QY 505 FQFQALQQAARH-----EGPLKCDISNSTEAGOKLFWMLRGKSEPTWLLAENVGA 558
 DB 702 FOIYGLCRESGQYVGPDRPKRPLHQCDIYRQPAAGNLIKFLMSKAGASQAPVLEETLRE 761
 QY 559 KNNMVRPLANTFEPLFTWLKDON--KNSFGWSDMSDPYADQSK 601
 DB 762 GRIDGTALREYFAFLLEMLRQENLRITMYGWMND-GDYCKRSIE 805

RESULT 17
 Q9VJVI PRELIMINARY; PRT; 792 AA.

AC 09VJVI:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CG17988 protein.
 GN ANCE-3 OR CG17988.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Abell J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA DePallos B., Delcher S., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Duthin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fjosek A., Gong F., Gorrell J.H., Gu Z., Guan P., Glasser K.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jolani M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Pelazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheibel F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003641; AAF53356.1;
 DR FlyBase: FBgn0032536; Ance-3.
 DR InterPro: IPR001348; Peptidase_M2.
 DR InterPro: IPR000130; Zn_Mpeptidase.
 DR Pfam: PF01401; Peptidase_M2; 1.
 DR PRINTS: PR00791; PEPDIPASEA.
 DR ProDom: PD004184; PEPDIPASEA.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR PROSITE: 792 AA; 92321 MW; 757EC47208902280 CRC64;

Query Match 21.7%; Score 841; DB 5; Length 792;
 Best Local Similarity 30.5%; Pred. No. 5e-53;
 Matches 196; Conservative 105; Mismatches 224; Indels 118; Gaps 14;

QY 30 WNYNTITEENVQNNMNNAGDKMSAFLEQGS--TLAQY-----PLQSIQNLTYK 76
 DB 158 YNNNNNVEYV--GINNNRNFENFLSNDRFNLDGRLRORYODRRYQGLTFLKRLTI 216
 QY 77 L-----QLQALQNGSSV 89

```

Db 217 LVSDQGSLECTNLAQOQRYVEFORITAQSKRINKOLIFRRRLRLQMLQSEVQNA 276
Qy 90 LSEDKSKRLNTLNTSTSTYSGKVCNPDPQEC-LILEPGLNEIMANSIDYERLMAE 148
Db 277 LPDLVDLRKYNLNLNEMFLVYSABICAYQQPFCODLHYIPQOLDIMAKSDMDLQHTWV 336
Qy 149 SWRSVGVKQRLPLVEEVVVLKNEAARNHEDYGDYRGVYENGVGYSGQLIEVY 208
Db 337 EYHRAAGKMDSDYEQOLDVQVEAYVNVNNGEYLAIE-----SGNRQDM 386
Qy 209 EHTFEIKPLVEHLHAYVRKALNNAY-PSYISPIGCLPAHLGDMGRTNLSYLTVE 267
Db 387 DIVEWQIRPLVEGLHAYVRRLRDYVGGDRINRLAPISHLGDMVQGSNVLIDILIPY 446
Qy 268 GQKPNIDVTAMDQANDQRIEKEAEKFFVSGLPMQTFGNSMLDTPGVOKAYCH 327
Db 447 PRRKRLIDVTPRVVEGGYPPQLMFLQAEFFTSIMSVAGPEFRKNSLFEQPLD-RRLVCE 505
Qy 328 PPAWML-GKDPRLMCTKVTMDPLFAHHEMGHIQYDMAVAAQPFLLRNGANGEPHEAV 386
Db 506 PPAWPCNRHDFRVAICTDINORSLSYHHEMAHQYFLQRIHLPIFRNGANPAFHQAV 565
Qy 387 GEIMSLSATPKRHLKISGLSPDQENETENFLKQALTYGTLEPTMLEKRMV 446
Db 566 GDAIGLSVTPRHLQTLGLQRLSIDE-SSYDINYLFWALDKVAFLEFALSLOWRVDV 624
Qy 447 KEEIFRDOMAKKMKREIYGVVPEVPHDEYCDPASLFHVSNDYSFIRYTYTLXQFO 506
Db 625 SGNAKRRYMNCHYMN-----LRFESTVLAQFO 651
Qy 507 FOEALQQAARH-----EGPLHKDISNSTEAGOKLFNMLRIGKSEPTLALENVGAKN 560
Db 652 ITRGICRESGQYVGPDPKRPFLHCIDYQRPAGNLIKTLMKSGASQPMQEVLEETLEGR 711
Qy 561 MNRVPLNFEPLFTWLDQN--KNSFGMSDMSPYVDOSIK 601
Db 712 LQGTALREYFAPLELWLRQENLRTEYVGMND-GDYCKRSIE 753

RESULT 18
Qy 99D836 PRELIMINARY; PRT; 249 AA.
Db 99D836;
AC 09D836;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE 2010305105Rik protein.
GN 2010305105Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid-10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-SMALL INTESTINE;
RX MEDLINE-21063660; Pubmed-11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arai K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana A. I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giusti C., King B., Kochiva H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh N.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momhaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata K., Storch K.-F.,
RA Suzuki H., Togo-Oka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

```

```

RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK008530; BAB5723.1; -.
DR MEROPS; M02.006; -.
DR MGI; MGI:1917256; 2010305105Rik.
DR InterPro; IPR001548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
SQ SEQUENCE 249 AA; 28379 MW; 19372B2878AAE921 CRC64;

```

Query Match 19.7%; Score 762; DB 11; Length 249;

Best Local Similarity 76.0%; Pred. No. 5.9e-48;

Matches 142; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

```

Qy 539 MLRLKSEPTLALENVGAKNNVRPLNVEFEPLFTWLDKQNSFVGNSTDMSPYADQ 598
Db 1 MSLGNSSEPTKALENVGAKNNVQKPLNVEFGLFQDLKEQNRNSFVGNSTDMSPYADQ 60
Qy 599 SIKVRSISKSLGDKATKENDNEMTLFRSSVAYAMRYFLKVKNNOMILGGEQVRANK 658
Db 61 SIKVRSISKSLGANAAYEMTNEMFLFRSSVAYAMRYFSITKNQVPLFEDVRSYDK 120
Qy 659 PRISFNFVTAPKRVSDIIPREVEKAIKMSRSRINDAFRLNDSLEFLGIQPTLGPNO 718
Db 121 PAVSYFFVTSPQNVSDVIRPSEVDALIRMSGRINQVGLNDSLEFLGIHPTLPYQ 180
Qy 719 PP 720
Db 181 PP 182

```

RESULT 19

Q18581 PRELIMINARY; PRT; 907 AA.

```

AC 018581;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Hypothetical 101.1 kDa protein.
GN C42D8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid-6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; Pubmed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Hallsworth K.;
RA "The sequence of C. elegans cosmid C42D8.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Waterston R.;
RA "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U56966; AAA98719.1; -.
DR InterPro; IPR001548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
DR ProDom; PD004184; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 907 AA; 101086 MW; F9B8B5C0F9BC5MA3 CRC64;

```

Query Match 16.5%; Score 638.5; DB 5; Length 907;

Best Local Similarity 27.0%; Pred. No. 4.2e-38;

[illegible]

	RESULT	20
ID	09V520	
AC	PRELIMINARY;	PRT; 661 AA.
DT	01-MAY-2000 (TrEMBLrel. 13,	Created)
DT	01-MAY-2000 (TrEMBLrel. 13,	Last sequence update)
DE	CG8196 protein.	
GN	ANCE-4 OR CG8196.	
OS	Drosophila melanogaster (Fruit fly)	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.	
CC	Ephydroidae; Drosophilidae; Sophophila.	
NCBI_TaxID=7227;	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-BERKELEY;	
RC	MEDLINE-20196006; PubMed-10731132;	
RX	Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA	Ananthes P.O., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,	
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Peiffer B.D.,	
RA	Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA	Abrill J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck R.J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Padlos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslcr A., Gargielan A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyram C.,
RA Jalali M., Kalush E., Kratt C., Krevitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Krevitz S., Kulp D., Lai Z.,
RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Munkulov G., Milshina N.V., Mobbary C., Morris J., Mostreil A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
RA Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenklatnos I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A., Sun E.,
RA Wang Z.-Y., Wasserman D.A., Weltscho G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.,
RT *The genome sequence of *Drosophila melanogaster*.*
RL Science 287:2185-2195(2000).
DR EMBL: AE003835 :AAR59002.1; -
DR FlyBase: FBgn003356; Ance-4.
DR InterPro: IPR001348; Peptidase_M2.
DR Pfam: PF01401; Peptidase_M2.1.
DR PRINTS: PR00791; PEPDIPTASEA.
DR PRODOM: PD004184; PEPDIPTASEA_M2.1.
SQ SEQUENCE 661 AA: 76218 MW: 20CB768AC607306E CRC64;

Query Match	12.7%	Score 491	DB 5	Length 661
Best Local Similarity	23.5%	Pred. No. 1.7e-27		
Matches 163	Conservative 133	Mismatches 298	Indels 100	Gaps 25

```

Dd 1 SYLEBARKFLKFNHEADLFYO--SLSASNYNNITEE-----AVOANNNAQDKWS 52
Dd 20 ANIDBAKTFV-----NQSDRYRRFNEIAETYSANNEEDDALESKLNNVRAETELV 75
QY 53 AFLEKESTIAOMPTLOEIONLTVKLOLALOONGS--SYLEBKSRLNTINTMSTYST 111
Dd 76 SISRKAAT-----YDIDRIRSPOTKMALDELFRAGDLFVLGDYFSSQOMNIALQTL-ST 130
QY 112 GFVCPN-----DNPOCILLEGLNEMANSIDYNERLMAWMSRSEVQKOLRP 160
Dd 131 DKDIEPYLGAGMNPEDDSP--LAFYFDIOKIVQSSKDQDELKYWMAMD----- 179
QY 161 LYEELVVLKNEH--APANITYEDYGDYWG-----DYEVNGVDYDYRSQGLIDVEHPREI 215
Dd 180 -----KNQWMASLNFTYTVQSYORAKILLEVPHKL--WYRYSQEMIQOMQOATEL 230
QY 216 KPLYEHLHAYRAKILMNAVPSYISPI--GCLPAHLISGMWGFRT--NLISLYVEFGKPN 272
Dd 231 RPAVOQLHAFVROELHKKYSDDVYNNLNGRPIDHILFPOOVLROQAWAGSILIDYPRQOLPE 290
QY 273 IDVDAMDQMDAQRIFEAEKEKFSVSGLEPNMTOGFWMSMLTDPQNVQKACHPYAMD 332
Dd 291 FDEYVSH-----TAKTAVNSEENYTTILGEPFLSAEFHKNQKLEPNQDSDNDCCRISIFD 346
QY 333 LKQKDFRILMCTKYTMDFTLHNEHEMHIQYDAAVYAOPPLLRNGANEGCHEHVEGSETMSI 392
Dd 347 LTPHYV--LMCEKSEVSRKLMQYHSHMARVYVYAAOKSHLSBPYRYKAVNLEF--AVGSAVYL 403
QY 393 SAAPRILKISGLSPFOEDNETENIFLKOALLTVGTLFPTMYMLEKRWMYFKEEIPK 452
Dd 404 SASPSAHLTGRRILAKEVLE--PALMSRILFRMAIHITLISLIPYVHTKYMHLNDLNDTYDM 461

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 17:58:24 ; Search time 4974 Seconds

(without alignments)
4212.710 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Sequence: 3869
1 STRIEGAKFTLDFKFNHEAD.....DNSLEFLGIGIPLGPPNP 720

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-O/cgn2.1/USPTO.spool/US09978385/runat_21052003_154419_19834/app_query.fasta_1.903
-DB=Genembl -GFWT=fastp -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -STRAP=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20 -MODE=LOCAL
-OUTPRT=Pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385.ecgn_1.1.3284 -runat_21052003_154419_19834 -NCPU=6 -ICPU=3
-NO_MAP -LARGOQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :

Genembl.*

1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ot:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2415	6 AR135178	AR135178 Sequence
2	3869	100.0	2415	6 AX418984	AX418984 Sequence
3	3869	100.0	2415	6 E43987	E43987 ACE-analogo
4	3869	100.0	2418	6 E39033	E39033 MPROT15 pol
5	3869	100.0	2599	6 E43988	E43988 ACE-analogo
6	3869	100.0	3325	6 AF291820	AF291820 Homo sapi
7	3869	100.0	3334	6 AX047758	AX047758 Sequence
8	3869	100.0	3341	6 HSM800880	AL110224 Homo sapi
9	3869	100.0	3396	6 AR135177	AR135177 Sequence
10	3869	100.0	3396	6 AX418982	AX418982 Sequence
11	3869	100.0	3405	6 AX412554	AX412554 Homo sapi
12	3869	100.0	3405	6 AB046569	AB046569 Homo sapi
13	3866	99.9	2599	6 AB046569	AB046569 Homo sapi
14	3720	96.1	3732	6 E39034	E39034 MPROT15 pol
15	3366.5	87.0	2262	6 E39034	E39034 MPROT15 pol
16	3251	84.0	2638	6 AX047762	AX047762 Sequence
17	3251	84.0	2739	10 BC026801	BC026801 Mus muscu
18	3251	84.0	2760	10 AB053181	AB053181 Mus muscu
19	3233	83.6	2638	6 AX047765	AX047765 Sequence
20	3175	82.1	2415	6 AX047760	AX047760 Sequence
21	2645	68.4	2415	6 AX047764	AX047764 Sequence
22	1844	47.7	1993	10 AB053182	AB053182 Mus muscu
23	1345.5	34.8	4879	9 BC036375	BC036375 Homo sapi
24	1337	34.6	2473	9 HUMRACER	M2657 Human testl
25	1337	34.6	2477	9 HUMRACER	X16295 Human testl
26	1337	34.6	2478	6 A31567	A31567 H.sapiens t
27	1337	34.6	2478	6 AR037213	AR037213 Sequence
28	1337	34.6	4020	6 AR137383	AR137383 Sequence
29	1337	34.6	4020	9 HUMRACER	J04144 Human angio
30	1337	34.6	4022	6 AX429555	AX429555 Sequence
31	1337	34.6	4024	6 A00914	A00914 H.sapiens g
32	1334	34.5	2418	10 MUSACER	M55333 Mouse testl
33	1334	34.5	3681	10 MUSACER	BC034367 Mus muscu
34	1334	34.5	3813	10 MUSACER	J04946 Mouse angio
35	1334	34.5	4563	6 AX147531	AX147531 Sequence
36	1334	34.5	4564	10 MUSACER	AX147531 Sequence
37	1334	34.5	4694	10 MUSACER	J04947 Mouse angio
38	1315	34.0	2551	9 HUMRACER	M2981 Human abert
39	1312	33.9	4050	5 CHKACE1	L40175 Gallus gall
40	1310	33.9	4014	10 AF201331	AF201331 Rattus no
41	1310	33.9	4014	10 AF201332	AF201332 Rattus no
42	1310	33.9	4014	10 RND03708	RND03708 Rattus norv
43	1303	33.9	4142	10 RND03734	U03734 Rattus norv
44	1303	33.7	2490	9 HUMRACER	M2658 Human testl
45	1283	33.2	2409	4 RABRACA	J05041 Rabbit angio

RESULT 1

ALIGNMENTS

QY 581 AsnlysnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 Db 1795 AACAGAAATCTTTTGGATGGATGAGTACCGACTGGATCCATATGACGACCAAGCATC 1854
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrAsnAspAsn 620
 Db 1855 AAGTAGAGTAAAGCTTAAATCAGCTCTTGGAGATTAAGCATATGAAATGGAACGACAAAT 1914
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuVal 640
 Db 1915 GAATGTACCTGTCGATCATCTGTGCATATGCTATGAGGCACTCTTTTAAAGCTA 1974
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlyAspValArgValAlaAsnLeuLysProArg 660
 Db 1975 AAAATACAGATATCTCTTTTGGAGAGAGATGTCGAGCGCTAAATTTGAAACCAAGA 2034
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2035 ATCTCTTAAATTTCTTGTCTACCTGACCTTAAATATGCTGATATCATCTTCTAGAACT 2094
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2095 GAAGTTGAAAGGCCATACAGATGTCGAGCGCTATCATGATGCTTCCGCTGAAT 2154
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2155 GACACACACCTAGAGTTCTGCGGATACAGCAACACTTGGACCTCTTAACGACGCCCT 2214
 RESULT 2
 AX418984 2415 bp DNA linear PAT 18-JUN-2002
 LOCUS AX418984
 DEFINITION Sequence 3 from Patent WO0212471.
 ACCESSION AX418984
 VERSION AX418984.1 GI:21523783
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Acton, S., Robison, K.E. and Hsieh, F.Y.
 TITLE Angiotensin converting enzyme homolog and uses therefor
 JOURNAL Patent: WO 0212471-A 3, 14-FEB-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..2415
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 743 a 483 c 555 g 634 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0
 Score: 3869.00 Length: 2415
 Percent Similarity: 100.00% Matches: 720
 Best Local Similarity: 100.00% Conservative: 0
 Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AX418984 (1-2415)
 QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 55 TCCACCATTTGAGAACAGCCAGCATTTTGGACAAGTTTAAACCGAAGCCGAAGAC 114
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluLysAsn 40
 Db 115 CTGTTCTATCAAAAGTCTCTTCTTGGATTAATTAACCAATATTTACTGAAGAAAT 174
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 Db 175 GTCCAAAACATGATATATGCTGGGACAAATGCTCTCTTTTAAAGAACAGTCCACA 234

QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 235 CTTGCCCAATATGATCATCTACACAGAAATTCAGAAATCTCACAGTCAACCTTCAGCTGAC 294
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 295 GCTCTTACCAAAATGGCTTCTCAGTGTCTGAGAAAGAACAGCAACGGTTGACACA 354
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 355 ATTCATTAATACATGATGACCCATCTACAGTACTGGAAAAAGTTTGAACCCAGATATCA 414
 QY 121 GlnGluCysLeuLeuLeuGluLeuProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 Db 415 CAAGAAATGCTTATCTTCAACCAAGCTTGAATGAATTAATGCCAAGATTAAGCTTAC 474
 QY 141 AsnGluArgLeuThrAlaTyrGlnSerTyrArgSerGluValGlyLysGlnLeuArgPro 160
 Db 475 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTGGGCAAGCAGTGAAGCCA 534
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 535 TTATATGAAAGATGATGCTGCTTGAATAATGAGATGCCAAGCAATTAATGAGAGAC 594
 QY 181 TyrGluAspTyrTyrPheArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 595 TATGGGATTAATGAGAGAGAGACTAATGAATGAGTAATGGGTAGATGCTATGACTACAC 654
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlu 220
 Db 655 CGCGGCGAGTGTGATGAAGATGAGAACATACCTTTGAGAGATTAATACCATTAATATGA 714
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 715 CATCTTATGACCTTATGAGAGGAGCAAGTTGATTAATGCTTATCTTCTATATGAGTCA 774
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 Db 775 ATTGAGATGCTCTCCCTGCTCATTTGCTGGAATATGGGATGATTTGGCAAAATGTG 834
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 835 TACTCTTTGACAGTCTCTTCTTGGACAGAAACCAATGATGATTAATGATGATGATG 894
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 895 GACCAAGGCTGGGATGACAGAAATATCAAGAGGCGGAAAGTTTGTATCTGCT 954
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 955 GGTCTTCTAATATGATCTCAAGATTCGGAATTCATGCTAAGCAGCCAGGAAT 1014
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 Db 1015 GTTCAGAAAGAGTGTGCGATCCACAGCTTGGAGCTGGGGAAGGCGATCTAGCATC 1074
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 Db 1075 CTTATGTGCACAAAGTACAAATGACATGCGATTCCTGACAGCTCATGAGATGGGCAT 1134
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuMetArgAsnGlyAlaAsnGlu 380
 Db 1135 ATCCAGTATGATATGATGATGATGCTGACACACCTTTCTGCTAAAGAAATGAGCAATGA 1194
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1195 GATTCATCAATGAAGCTTGGGGAATATCATGCTTCTGAGGACACCTTAAGCATTTA 1254
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluLysAsnGluThrGluIleAsnPhe 420
 Db 1255 AAATCATTTGCTCTTCTGCTACCCGATTTTCAAGAAACATATAACAGAAATTAACCTTC 1314
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440

```

|||||
1315 CTGCTCAACAGCAGCTACGATTTGGGAGCTGCTCCATTACTTTCATGATGAGAG 1374
441 TTPATGTTPMetValPheLeuGlyGluLeuProLysAspGlnTrpMetLysLysTyrPro 460
1375 TGGAGCTGATGCTGCTTTTAAAGGGAAATCCCAAGACAGCTGAGTGAAGAAAGTGTGG 1434
461 GtMetLysAspGluLeuValGlyValValGluProValProHisAspGluTrpTyrCys 480
1435 GAGATGAGCGAGATGATGGGGTGGAGCTGGCCCATGATGATGAACATGATGCT 1494
481 ASPProLysSerLeuPheHisValSerAspTyrSerPheLeuArgTyrTyrTrpArg 500
1495 GACCCCGCATCTCTGTCATGCTTCTATGATGATGATGATGATGATGATGATGATGATG 1554
501 ThrLeuTyrGlnPheGlnPheGlnGluLeuValGluValGluValGluValGluValGlu 520
1555 ACCCTTACCAATTCAGATTCAGAGACAGCTTTCAGAGCAGCTTAAACATGAGAGCCCT 1614
521 LeuHisLysCysAspLysSerAspSerThrGluValGluValGluValGluValGluVal 540
1615 CTGACAAATGTCATCTCAATCAATCTCAATCAATCTCAATCAATCTCAATCAATCTG 1674
541 ArgLeuGlyLysSerGlnProTrpPheLeuValGluValGluValGluValGluValGlu 560
1675 AGCGTTGAAATACAGAACCTTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 1734
561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheTrpLeuLysAspGln 580
1735 ATGATGTAAGCCACGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1794
581 AsnLysAsnSerPheValGlyTyrPheTrpAspTyrPheTrpAspTyrPheTrpAspTyr 600
1795 AACAGATTTCTTTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1854
601 LysValArgLysSerLeuLysSerAlaLeuGlyLysPheValGlyTyrPheTrpAspAsn 620
1855 AAGTGAAGATTAACCTTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1914
621 GtMetTyrLeuPheAspSerSerValAlaTyrTrpMetLysGlnTyrPheLeuValGlu 640
1915 GAAATGTACCTGCTCCGATCATCTGTCATATCTGATGAGGAGCTTCTTTTAAAGTA 1974
641 LysAsnGlnMetLeuPheGlyGluGluValAlaTyrValAlaAlaAsnLeuLysProArg 660
1975 AAAATTCAGATGATCTTTTGGGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAG 2034
661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
2035 ATCTCCCTTAATTTCTTGTCTCTGCTCACTCACTCACTCACTCACTCACTCACTCACT 2094
681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
2095 GAGTTGAAAGAGCCATCAGATGTCCTCGGAGCCCTATCAATGATGCTTCCGCTGAAT 2154
701 AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
2155 GACAAACAGCTAGATTTCTGGGAGATACAGCCAACTTGACCTTCAACAGCCCTT 2214

```

RESULT 3
 E43987
 LOCUS ACE-analogous gene. 2415 bp DNA linear PAT 31-JAN-2002
 DEFINITION E43987
 ACCESSION E43987.1 GI:18629190
 VERSION JP 2001046072-A/1.
 KEYWORDS
 SOURCE unidentified.
 ORGANISM
 unclassified.
 1 (bases 1 to 2415)
 SUGANO, S. and KOMATSU, T.
 ACE-analogous gene
 Patent: JP 2001046072-A 1 20-FEB-2001;
 JOURNAL

COMMENT

OTSUKA PHARMACEUT CO LTD

OS Unknown
 PN JP 2001046072-A/1
 PD 20-FEB-2001
 PE 06-AUG-1999 JP 1999223892

PR SIMIO SUGANO, TAKAMI KOMATSU
 PI C12N15/09, A61K31/00, A61K31/77088, A61K38/00, A61K38/55,
 PC A61K39/395, PC A61K39/395,
 PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC
 C12N1/19,
 PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC
 A61K37/64,
 CC C12N5/00

CC key location/Qualifiers
 FH key 1.2415
 FT source /organism="unknown"
 FT Location/Qualifiers

FEATURES
 source
 1.2415
 /organism="unidentified"
 /db_xref="taxon:32644"

BASE COUNT 743 a 484 c 554 g 634 t

ORIGIN
 Alignment Scores:
 Pred. NO.: 0 Length: 2415
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x E43987 (1-2415)

```

1 SerThrIleGluGluGlnAlaLysTrpPheLeuAspLysPheAsnHisGluAlaGluAsp 20
55 TCACCAATGAGGAAAGCCAGCAAGACATTTTGGAGCAAGTTTAAACACGACCAAGCCAGAGC 114
21 LeuPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
115 CTGTTCTATCAAGTTCATCTGCTTCTTGCAATTTATACACCAATTTACTGAGAGAT 174
41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrPheSerAlaPheLeuLysGluGlnSerThr 60
175 GTCCAAACATGAAATTAATGCTGGGAGCAAAATGCTCTTTTAAAGAGCAAGTCCACA 234
61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
235 CTGGCCCAATGATATCCATCAAGAAATTCAGAAATTCACAGTCAAGTTCAGTGCAG 294
81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysAlaGluAsnThr 100
295 GCTCTTCACCAAAATGGGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 354
101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
355 ATCTTAATATCAATGATGAGCAGCATCTACAGTCTGAGAAAGTTTGAACCAAGTAAATCA 414
121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
415 CAAGATGCTTATTTACTTAAACAGGTTTGAATGAATGAATGAATGAATGAATGAATGAAT 474
141 AsnGluArgLeuTrpAlaTrpGluSerTrpAspSerGluValGlyLysGlnLeuArgPro 160
475 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATGAGTGGAGGAGGAGGAGGAGGAGGAG 534
161 LeuTyrGlnGluLysValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
535 TTATATGAGAGATGATGCTGCTTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 594
181 TyrGlyAspTyrTrpArgLysPyrGluValAsnGlyValAspGlyTyrAspTyrSer 200

```

Db 595 TATGGGATTTATGAGAGGAGACTATGAAAGTAAATGGGTAGATGCTATGACTACAGC 654
 QY 201 ArgGlyGlnLeuIleGluAspValGluIsthrPheGluGluIleLysProLeuTyrGlu 220
 Db 655 CGCGGCGCATGATGATGAGATGTGGAACATACCTTTGAGAGATTAACCACTTATATGAA 714
 QY 221 HistHHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 715 CATCTTCATGCTTATGTCAGGCAAGTGTGATGATGATGATGATGATGATGATGATGAT 774
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGluIleAspMetThrPheGluTyrPheThrLeu 260
 Db 775 ATGTGATGCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 835 TACCTCTTGGACAGTTCCTTGGACAGCAACCAACATAGATTTACTGATGCAATGCTG 894
 QY 281 AspGlnAlaThrAspAlaGlnArgIlePheLysGluAlaLysPhePheValSerVal 300
 Db 895 GACCAAGGCTGGATGCAAGCAAGATATTCAGAGAGCCGAGAGAGTTCCTTGTATCTGTT 954
 QY 301 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 955 GGTCTCTTATATGACTCAAGATTCCTGGAAATTCCTGCTACGAGCCAGGAAAT 1014
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 Db 1015 GTTATGAAAGCAGCTGCTCCATCCACAGCTTGGAGAGCCGAGAGAGTTCCTGCTGAG 1074
 QY 341 LeuMetCysThrLysValThrMetAspPhePheLeuThrAlaHisGluMetGlyHis 360
 Db 1075 CTATATGTCACAAAGGAGCAATGAGAGAGCTTCTACAGCTCATCATCATCATCATCAT 1134
 QY 361 IleGlnTyrAspMetAlaTyrAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1135 ATCCAGATATATGATATGATATGCTGCACAACTTTCTCTTAAGAAATGAGACTATGAA 1194
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1195 GGAATCCATGAACTGTGGGAAATCATGCTCCTTCTGCGAGCCACACCTTAAGACTTAA 1254
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 Db 1255 AAATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 Db 1315 CTGCTCAACCAACGACTCAGATGTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
 QY 441 TrpArgTyrMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTyrTrp 460
 Db 1375 TGGAGGTGATGCTGCTTAAAGGGGAATCCCAAGACAGCATGATTAAGAAAGGTGG 1434
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 Db 1435 GAGATGAACGAGAGATGTTGGGTGGGAACTGTGCCCATGATGAAACAACATGCT 1494
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 Db 1495 GACCCCGCATCTCTGCTCATGTTCTTAATGATTAATGATTAATGATTAATGATTAATG 1554
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 Db 1555 ACCCTTTACCAATTCAGTTTCAAGAACACCTTGTCAAGAGCTTAAACATGAGAGCCCT 1614
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeu 540
 Db 1615 CTGCAACAATGATGATCTCAAACTCTACAGAGCTGAGCAAGAACTGTTCAATATGCTG 1674
 QY 541 ArgLeuGlyLysSerGluProTyrThrLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
 Db 1675 AGGCTTGGAAATCAGAACCTGAGCCCTGATGGAATGTTGAGAGAGCAAGAAC 1734

QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGln 580
 Db 1735 ATGATGTAAGGCGACCTCTCAACACTTACAGCCCTTATTTACTTGGCTGAAAGACCG 1794
 QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 Db 1795 AACAAAGAAATCTTTTGGTGGATGAGTACGACAGCTGAGTCCATATGACCAAAACGATC 1854
 QY 601 LysValAlaGlyIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsn 620
 Db 1855 AAAGTGAAGATTAAGCTTAATATCACTCTTGGAGATTAACCAATGAAAGACGCAAT 1914
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysVal 640
 Db 1915 GAAATGATACCTGTTCCCATCATCTGTCATATGCTATGAGGAGGAGTCTTTAAAGTA 1974
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 Db 1975 AAAATATGATGATCTTTTGGGAGAGAGATGTGAGTGTGCTAATTTGAAACCAAGA 2034
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2035 ATCTCTTATATTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094
 QY 681 GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2095 GAACTTGAAGAGGCGATCAGATGATGCTCCGAGAGCGGTATCAATGATGCTTCCGCTGAA 2154
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2155 GACAAAGAGCTAGATGTTCTGGGAGATACAGCCAACTTGGAGCTCTTACCAAGCCCT 2214
 RESULT 4
 E39033
 LOCUS E39033 2418 bp DNA linear PAT 18-JUN-2001
 DEFINITION MPROT15 polypeptide and MPROT15 polynucleotide.
 ACCESSION E39033
 VERSION E39033.1 GI:13017695
 KEYWORDS JP 199318472-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2418)
 AUTHORS Christopher, D.S. and Nicola, B.
 TITLE MPROT15 polypeptide and MPROT15 polynucleotide
 JOURNAL Patent: JP 199318472-A 1 24-NOV-1999;
 SMITHKLINE BEECHAM CORP PUBLIC LTD CO
 OS Homo sapiens (human)
 PN JP 199318472-A/1
 PD 24-NOV-1999
 PF 22-JAN-1999 JP 1999014949
 PR 13-MAY-1998 GB 9810373; 2.18-AUG-1998 GB 9818009; 4 PI
 CHRISTOPHER D SAZAN, NICOLA BAGEBU
 PC C12N15/09, A61K31/00, A61K31/00, A61K31/00, A61K31/70,
 PC A61K38/00,
 PC A61K39/395, A61K45/00, A61K48/00, C07K14/47, C12P21/02, C12Q1/68,
 PC GOIN33/15,
 PC GOIN33/50, GOIN33/50, GOIN33/56, GOIN33/68, C12P21/02,
 PC C12R1/91, PC C12N15/00,
 PC A61K37/02
 CC CC
 FH Key
 FT source 1. 2418
 FT Location/Qualifiers
 1. 2418
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 744 a 484 c 555 g 635 t
 ORIGIN

QY 681 GUAValGluLysAlaIleArgMetSerArgSerA-gliLeuAspAlaPheArgLeuAsn 700
 DB 2095 GAAGTTGAAAGAGCCATTCAGATGTCGCCGACCCCTATCAATGATGCTTCCGCTGAT 2154
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2155 GACACAGAGCTTGAAGTTCTGTGGGATACAGCCACACTTGGACCTCTTACACAGCCCT 2214

RESULT 5
 E43988
 LOCUS E43988 2599 bp DNA linear PAT 31-JAN-2002
 DEFINITION ACE-analogous gene.
 ACCESSION E43988.1 GI:18629191
 VERSION JP 2001046072-A/2.
 KEYWORDS unclassified.
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 2599)
 AUTHORS Sugano,S. and Komatsu,T.
 TITLE ACE-analogous gene.
 JOURNAL Patent: JP 2001046072-A 2 20-FEB-2001;
 OTSUKA PHARMACEUT CO LTD
 COMMENT OS Unknown
 PN JP 2001046072-A/2
 PD 20-FEB-2001
 PF 06-AUG-1999 JP 1999223892

PI SUMIO SUGANO, TAKAMI KOMATSU
 PC C12N15/09, A61K31/00, A61K31/7088, A61K38/00, A61K38/55,
 A61K39/395, PC A61K39/395,
 PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC
 C12N1/19,
 PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC
 A61K37/64,
 PC C12N5/00
 CC
 FH
 FT Key Location/Qualifiers
 CDS Location/Qualifiers
 1..2599
 /organism="unclassified"
 /db_xref="taxon:32644"

BASE COUNT 807 a 503 c 591 g 698 t

ORIGIN

Alignment Scores:
 Pred. No.: 0
 Score: 3869.00 Length: 2599
 Percent Similarity: 100.00% Matches: 720
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 100.00% Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x E43988 (1-2599)

QY 1 SerThrIleGluGlnAlaIleArgMetSerArgSerA-gliLeuAspAlaPheArgLeuAsn 20
 DB 109 TCCACCATTCAGAGAGCCCAAGCATTTTGGACAGATTAAACAGAGAGAGAGAGAGAGAG 168
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheThrValAsnThrAsnIleThrGluGlnAsn 40
 DB 169 CAGTTCTATCAAGAGTTCCTTCTGTGGAATTAACACCAATATTAATCAAGAGAGAGAG 228
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 229 GTCCAAACATGAAATGCTGGGACAAATGGTCTGCTTTTAAAGAGAGAGAGAGAGAG 288
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 289 CTTCGCCAATATGATCACTACAAAGAAATTCAGATTCACAGCTCAAGCTTCAGCTGACG 348
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100

DB 349 GCTCTTCAG 408
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 409 ATCTCAATATCAATGAG 468
 QY 121 GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 DB 469 CAGAGATGCTTATTCCTTGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTyrPheArgSerGluValGlyLysGlnLeuAspPro 160
 DB 529 AATGAG 588
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 589 TTATATGAAGAGATATGCTGCTGAAAATGAGATGGCAGAGAGAGAGAGAGAGAGAGAG 648
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 649 TATGGGAGATATGAG 708
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlu 220
 DB 709 CGCGGAG 768
 QY 221 HisLeuHisAlaTyrValAlaGlyAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 769 CATCTTCATGCTATGAG 828
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGluAspMetTrpGlyArgThrPheThrAsnLeu 260
 DB 829 ATGTGAG 888
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 889 TACTCTTTCAG 948
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPheValSerVal 300
 DB 949 GACAG 1008
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1009 GGTCTTCCTAATATGAG 1068
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTle 340
 DB 1069 GTTCAG 1128
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetCysHis 360
 DB 1129 CTTATGTGAG 1188
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGlu 380
 DB 1189 ATCCAGATGATGATGAG 1248
 QY 381 GlyPheHisGluAlaValGluIleLeuSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1249 GATTCATCAAG 1308
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1309 AATATCATGAG 1368
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyLysLeuProPheThrTyrMetLeuGlnLys 440
 DB 1369 CTTCCTCAAG 1428
 QY 441 TrpArgTrpMetValPheLysGlyLysLeuProLysAspGlnTrpMetLysLysTrpTrp 460

Db	1429	TCGAGGTGATGGCTCTTTAAAGGGGAATTC	CCCAAGACAGCATGTA	AAAAAGTGGTGG	1488	
Oy	461	GIUWETLYSRAGGUJLIEVALGIYVALJG	UPROVALPROHISASPIGUTHRYCYLS		480	
Db	1489	GAGATGAAAGCAGAGATAGTGGGGTGGTGGAACTG	TGGCCCATATGAAACATCTGT		1548	
Oy	481	ASPRPLASERLEUPHEHISVALSERASNA	SPTYSERPHEHLEAGTYRYTTRATG		500	
Db	1549	GACCCCGCATCTGCTCCATGCTTTCTTAATGAT	TACTCATTCATTTGATATTTACACAGG		1608	
Oy	501	THIRLEUTYRGINPHEGINPHEGINGLUALALEUC	SGINLAALALYSHISGLUGLYPRO		520	
Db	1609	ACCCTTACCAATTCAGTTCACAAAGACACTT	GTTCACACACTAAACATGGAAGGCCCT		1668	
Oy	521	LEUTHISYCSAAPPILIESERASNSERTHGLUAL	AGLYGINLYSLEUPHEASMETLEU		540	
Db	1669	CTGCACAAAGTGTACATCTCAACTCTACAGAAC	CTGCAGAAACCTGTCANATATGCTG		1728	
Oy	541	ARGLEUGLYLYSSERGIUPROTPTPTHLEUALALEU	GLUASNAVALVALGIYALALYSASN		560	
Db	1729	AGCGTTGGAAATCAGAACCTCGACCTACAT	TGGAAATGTTGTAGGCGCAAAACAC		1788	
Oy	561	MELASNAVALARGPROLEULEAASNTYRPH	GLUPROLEUPHEHTRTPLEULYSASGLN		580	
Db	1789	ATGAATGTMAAGCCACCTGCTCACTACTT	TAGCCCTTATTTACTGCTGAAAGACCAG		1848	
Oy	581	ASNLVSASNSERPHEVALIGLYTRPSE	THRASPTPSEPROTYRALAASPGLNSERILE		600	
Db	1849	AACAAGATTCCTTTTGTGGGAGTGGAGTAC	CCAGATGCATATGCAGACCAAGCATC		1908	
Oy	601	LYSVALARGLIESERLEULYSERLALAEUL	LYASPYLSALATYRGUTRPAASAPSN		620	
Db	1909	AAAGTGAGAGATTAAGCTTAAATACAGCT	CTTGGAAGTAAAGCATATGATGAGACACAT		1968	
Oy	621	GLUMECTYRLEUPHEARGSE	SERVALAATYRALAMEATARGINTYRPH	LEULYSVAL	640	
Db	1969	GAAATGTACCTTCTCCGATCATCTTG	TCATATGCTATGAGGACGACTTTTAAAGTA		2028	
Oy	641	LYSAANGIMETILELEUPHEGLYGLUGL	USAPVALARGVALALAAASLEULYSPROARG		660	
Db	2029	AAAAATCAATATATCTTTTGGGAGGAGGAT	GTCGGATGCGGCTTAATTTGAAACCAAGA		2088	
Oy	661	ILIESERPHEASN	PHEPHEVALTHRLAAPROLYSAANVALSER	ASPDILELEPROARGTHR	680	
Db	2089	ATCTCTTTAATTTCTTTCTACTG	CACTAAANAATGCTCATATCATTCCTAGAACT		2148	
Oy	681	GIUVALGIULYSALALIEARGMET	SERATARGISERARGILIASNA	ASAPALAPHEARGLEUNSN	700	
Db	2149	GAAGTGAAAGGCCATCTCAGATGTC	CCGAGCCCGATCAATGATGCTTCCGCTGAT		2208	
Oy	701	ASPAASERLEUNG	INPHELEUGLYLIEGINPROTHLEUGLYPRO	PROASNGINPROPO	720	
Db	2209	GACACACAGCCTAGATTTCTGGGGAR	ACAGCCACACTTGGAACCTCTTAACCAAGCCCT		2268	
RESULT 6	AF291820		3325 bp	mRNA	linear	PRI 15-JUN-2007
LOCUS	AF291820					
DEFINITION						
ACCESSION	AF291820					
VERSION	AF291820.1					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						

REFERENCE	PUBMED	10969042
2 (bases 1 to 3325)		
AUTHORS	Donoghue,M., Woolf,B., Robison,K. and Acton,S.	
TITLE	Direct Submission	
JOURNAL	Submitted (01-AUG-2000) Cardiovascular Biology, Millennium Pharmaceuticals, Inc., 75 Sidney Street, Cambridge, MA 02139, USA	
FEATURES	Location/Qualifiers	
source	1. .3325	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="X"	
	21. .2438	
CDS	/codon_start=1	
	/product="ACE-related carboxypeptidase ACE2"	
	/protein_id="AAFY9721.1"	
	/db_xref="gi:9802433"	
	/translation="MSSSSNLLSLVAVTAOSTIEQAKFLDKFNHREADLFYQSSLSAWNNTITEENTYNNMNNMGLKSAFLKEOSTLQAMPLOEIONLVKQLQALQOLNKSYSLSDEKSKRLNTLNTMTSTYSGTKGNDNPOECLELPGENITMANSLDYNSRLMANSNHSSEYGRKOLRPLREYVYLKEMARANHEDYGDWRDYEYNGVDGYSRQGLLEDVHEPTEETKPLREHLHAIVYAKLMANYPSTISPIGLPAHLIGDMCKEFTMLNLSYLPVPGKGRPNIDVDAMVDAMDAQRIPEAEKFFVSGLEPMHOGRENSMLLDPEANVAKACHPTAMDLGKDFRIIMCTKYTMDFLVAHENGHIQYDAIAQOPFLIRNCAANGAEVGEINSLSATPKHLKSIGLSPDEQENREINFLKQALTIVGITPEYLMEEKRMKRVFGEIPEKDOOMKMKMEKREIVGEYVEPDEYICDPAISLFVNSDYFIETKYRTIRTIKORPOEALCOAAKHGRLPHDKDINSSTENAGOKLFNLRLGKSEPPYALLENVGAKKMANNRYPLNTPETPLMTKIDONKNSVPGASDPYASQIKRISIKSFMSLAGRKNVDIIIPTEWEKAIKMSRSKINDAFLNDNSLEFLGLOPTLGPPNQPPYFIMVIGVYMGVIVVGIILFTGIRDRKKKNKRSGENPYASIDISKENNPQPNVTDVQTSF"	
BASE COUNT	1022 a 637 c 747 g 919 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	0	Length: 3325
Score:	3869.00	Matches: 720
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0
US-09-978-385-2.COBY_19_738 (1-720) x AF291820 (1-3325)		
Oy	1 SerThrIleGluGluGlnAlaIatysThrPheLeuAspIysPheAsnHISgluaIaGluasp 20	
Db	75 TCCACACATTGGAGGAACGGCCAGACATTGTTGGACAAGTTAAACCCAGGCGAAGAC 134	
Oy	21 LeuPheTygIlnSerSerLeuAlaSerTyrPasnTyAsnThrAsnIleThrGluGluAsn 40	
Db	135 CTTGTTCTTCAAGTCACTGCTCTTGGAATTAATTAACCCAAATATTATCGAAGACAAAT 194	
Oy	41 ValGlnAsnMetLanAsnAlaGlyAspIysTyrPserAlaPheLeuIysGluGlnSerThr 60	
Db	195 GTCCAAAACGATGATTAATGCTGGGGACAAATGCTGCTTTTAAAGAAACACTCCACA 254	
Oy	61 LeuAlaGlnMetTyrrProLeuGlnGluIleGlnAsnLeuThrValIysLeuGlnLeuGln 80	
Db	255 CTTGGCCAAATGATATCCCTACAGAAATTCAGAAATCTCACGTCACGCTCAG 314	
Oy	81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspIysSerIysArgLeuAsnThr 100	
Db	315 GCTCTTCACACAAATGGGCTTTCAGTCTCTCGAAGACAAAGACCAACGGTTGACACA 374	
Oy	101 IleLeuAsnThrMetSerThrIleTyrrSerThrGlyIysValCysAsnProAspAsnPro 120	
Db	375 ATTCTTAATACAAATAGACACATCTACAGTACGAAAGATTGTATACCCAGATTATCCA 434	
Oy	121 GlnGluCysLeuLeuLeuGluProGlyIysAsnGluIleMetAlaAsnSerLeuAspTyr 140	
Db	435 CAGAGATGCTTTATCTTCTGGAACACAGGTTTGAAAGAAATATGGCCAAACAGTTTACATC 494	
Oy	141 AsnGluAsnLeuThrAlaTyrPcolUserTyrPargSerGlnValGlyIysGlnLeuAsnPro 160	


```

Db      495 AATGAGAGGCTCTGGGCTGGGAAAGCTGGAGATCTGGCTGGCAAGCAGCTAGAGGCCA 554
Qy      161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
Db      555 TTATATGAAAGATATGAGTCTTAAATAATAGATGGCAGAGCAAAATCATATAGAGGAC 614
Qy      181 TyrGlyAspTyrTrrPahGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
Db      615 TATGGGGAATTAATGGAGAGAGACTATGAAGTAATAGGGAGATGGCTATGACTACAGC 674
Qy      201 ArgGlyGluLeuLeuGluAspValGluHisThrPheGluGluLeuLysProLeuTyrGlu 220
Db      675 CGCGGCCAGTGTGATGAAAGATGGAACATACCTTTGAAGAGATTAACCATATATGAA 734
Qy      221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLysSerPro 240
Db      735 CATCTCATGCTCATGTGAGAGGCAAAAGTTGATGATCCATCTTCCTATACAGCCA 794
Qy      241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrrPheAsnLeu 260
Db      795 ATTTGATGCTCTCCCTGCTCATTTGCTGTGATATGTG3GGTATTTTGGCAAAATCTG 854
Qy      261 TyrSerLeuThrValProPheGlyGluLysProAsnLysAspValThrAspAlaMetVal 280
Db      855 TACTCTTGACAGTTCCTCTTGAGCAGAAACCAACATAGATGTACTGATGCAATGTG 914
Qy      281 AspGluAlaTrrPaspAlaGluArgGlyLeuPheLysGluValGluLysPhePheValSerVal 300
Db      915 GACCAAGCTCTGGGATGACAGAGAAATATCAAGAGGCCGGAAGTTCTTGTATCTGT 974
Qy      301 GlyLeuProAsnMetThrGluGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
Db      975 GGTCTTCCTAATATGACTCAAGATTCCTGGGAAATTCATGCTTAACGGCAGCAAGAAAT 1034
Qy      321 ValGluLysAlaValCysHisProThrAlaTrrPaspLeuGlyLysGlyAspPheArgGly 340
Db      1035 GTTCAGAAAGCAGTGTGCCATCCACAGCTGGGAGCTGGGAGGAGCCACTTCAGAGATC 1094
Qy      341 LeuMetCysThrTrrPaspAlaThrMetAspPheLeuThrAlaHisGluMetGlyHis 360
Db      1095 CTATATGGCAGAAAGGTGAAATGAGACACTTCTCAGACGCTCATCATGATGAGTGGGCAT 1154
Qy      361 IleGluTyrAspMetAlaTrrAlaAlaGluProPheLeuLeuArgAsnGluLysAsnGlu 380
Db      1155 ATCCAGATGATGTGCATATGCTGCACAACTTTTCTCTTAAGAAATGAGCTAATGAA 1214
Qy      381 GlyPheHisGluAlaValGlyGluLysMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db      1215 GGATTCATGAAAGCTGTGGGAAATCATGCTTTCGAGGCCACACTTAAGCATTTTA 1274
Qy      401 LysSerIleGlyLeuLeuSerProAspPheGluGluAspAsnGluThrGluLysAsnPro 420
Db      1275 AATTCATGTGCTTGTGTACCCGATTTTCAAGAGCAAAAGCAAAATTAACCTTC 1334
Qy      421 LeuLeuLysGluAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlyLys 440
Db      1335 CTGCTCAAAACAAGCACTCAGATTTGTGGACTGCGCCATTACTACATGTTAAGAGAG 1394
Qy      441 TrrPaspTrrMetValPheLysGlyGluLysProLysAspGluTrrMetLysLysTrrP 460
Db      1395 TGGAGTGGATGCTTTTAAAGGGGAAATTCCTCAAGCACTGATGAAAGAGTGGGG 1454
Qy      461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrLys 480
Db      1455 GAGATGAAAGCGAGATAGTGGGTGGTGGAACTGTGCCCCATGATGAAACATACGTGT 1514
Qy      481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrrPasp 500
Db      1515 GACCCCGCATCTCTTCATGCTTCTAATGATCTCATCATTCATTCGATATTACCAAGG 1574
Qy      501 ThrLeuTyrGluPheGluPheGluGluLysGluLysGluAlaLysHisGluGluPro 520

```

```

Db      1575 ACCCTTACCAATTCAGTTCATGAAGACACTTTGTCAGACGACTAAACATGAGCCCT 1634
Qy      521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGluLysLeuPheAsnMetLeu 540
Db      1635 CTGCACAAATGATACATCTCAAACTCATACAGAGAGCTGCAGAGAACTGTTCAATATCG 1694
Qy      541 ArgLeuGlyLysSerGluProTrrPheLeuAlaLeuGluLysValValGlyLysAsn 560
Db      1695 AGCCTTGAAATATCAACACCTTGAGCCCTGACCTGATGGAAATGTTCTAGAGCAAGAAC 1754
Qy      561 MetAsnValArgProLeuLeuAsnTrrPheGluProLeuPheThrTrrLeuLysAspGlu 580
Db      1755 ATGAATGTAAGGCACTGCTCAACTACTTTGAGCCCTTTTACCTGGCTGAAAGACAG 1814
Qy      581 AsnLysAsnSerPheValGlyTrrPheThrAspTrrPheProTrrAlaAspGluSerIle 600
Db      1815 AACAAATATCTTTTGTGGATGAGTACCGACTGAGTGCATATGACAGACCAAGCAATC 1874
Qy      601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrPaspAsn 620
Db      1875 AAAGTGAAGATTAAGCTTAATATCAGCTCTTGAGAGATTAAGCATATGAAATGAAAGCAAT 1934
Qy      621 GluMetTyrLeuPheArgSerSerValAlaTrrAlaMetArgGluTrrPheLeuLysVal 640
Db      1935 GAAATGTACTGCTTCGATCATCTGTGATATGATGATGAGGACGACTTTTAAAGTA 1994
Qy      641 LysAsnGluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db      1995 AAAATCATATATCTTTTGGAGAGGATGTCGAGAGTGCATTAATTTGAACCAAGA 2054
Qy      661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleGluProArgThr 680
Db      2055 ATCTCTTAAATTTCTTTGTCTACTGACCTAAATATGCTGATATCTTCCTGAGACT 2114
Qy      681 GluValGluLysAlaIleArgMetSerArgSerArgGlyLeuAsnAspAlaPheArgLeuAsn 700
Db      2115 GAAGTTGAAAGGCCATCAGATGTCCCGAGCCGTATCAATATGATGCTTCGCTGAT 2174
Qy      701 AspAsnSerLeuGluPheLeuGlyIleGluProThrLeuGlyProProAsnGluProPro 720
Db      2175 GACAAAGCTCATGATTTCTGGGGATACAGCCACACTTGGACCTCTCAACACCCCTT 2234

```

RESULT 7
AX047758
LOCUS AX047758 3334 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 1 from Patent WO0070032.
ACCESSION AX047758
VERSION AX047758.1 GI:11876765
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3334)
AUTHORS Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
TITLE Zacc2: a human metalloenzyme
JOURNAL Patent: WO 0070032-A 1 23-NOV-2000;
FEATURES ZymoGenetics, Inc. (US)
source Location/Qualifiers
1..3334
/organism="Homo sapiens"
/db_xref="taxon:9606"
35..2452
/feature="unnamed protein product"
/codon_start=1
/protein_id="GI:11876766"
/translation="MSSSSWLLSLVAVTAOSTIEQARTFLDKFNHEADLFYSS
LASNNVTNTTEENVQNNMNAAGKWSAFLEQSTLQWYPIQIOLIVYLQIALQALQ
NGSSVLSDESRLNTILNTMSTIYSGKVCNDPQECILPGLNEIYANSLDYN
RLWAMESRSEVQKRLPLIEEYVVLKENARHNHEDYDGRGYEVNGVNGVYYS
RGOLIEDEVETFEIKPLYLHAYVAKLMMNVPYSYISPIGCLPHILGDMGRWT
NLVSLVPEFGKRNIDVTLDAMVDQANDQAKIFKEAEKFPYSGVLPNMGFWNSMTT

DPGNOKAVCHPTAMDIGKDERILMCTKVTMDLFAHHEKHIOYMAVAPOEFL
 RKGANECHREANGELMSLSAATPRKIKSGLISPDOPDNTEINELIKKALITVGLT
 PETYMEKRMWVFEIPEKDWKMKBMKREIVGYEPHPDHTYCDPASTIEVSN
 DYSFIRYTRLYOFQFOEALCOAAKHGEPHKKCDIINSTEKQALFMKRLGSEPM
 TLAEENVGAKMNVRRPLINFEPLFTMLKQDNKNSFVWGSDWSPYADOSIKRISL
 KVALDGKAEKNDNEMYLEFRSSVAYAMQYELFKQKQILGSEEDVRVANKRISFN
 FVTAIPKRVNSDIIPRTVEKAIIRMSRIRNDAPFLNDSLEFLGIOPILGPNOPVS
 IWLIVGYMGIYVIGIYLLFTGIRDRKKKKKARSBNPASIIDSKEINNPGFQNT
 DDVOTSF*

BASE COUNT 1011 a 640 c 754 g 929 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 3334
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AX047758 (1-3334)

QY	1	SerThrIleGluGlnAlaIleValThrPheLeuAspLysPheAsnHisGluAlaGluAsp	20
DB	89	TCCACCATGAGGACAGGACGACATTTTGGACAAAGTTTAAACACGACGACGAGAGAC	148
QY	21	LeuPheThrGlnSerLeuAlaSerTyrPheAsnThrAsnIleThrGluGluAsn	40
DB	149	CTTTCTTATCAAGTTCACTTCTTCTTGGAAATTAACACCAATTTTCTGACAGAGAT	208
QY	41	ValGlnAsnMetAsnAlaGlyAspLysTyrPheAlaPheLeuGluGlnSerThr	60
DB	209	GTCGCAAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTGCACA	268
QY	61	LeuAlaGlnMetThrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln	80
DB	269	CTTGCCCAAAATGATCCACTACACAGAAATTCAGAAATCTCAGCTCAGCTCAGCGAG	328
QY	81	AlaLeuGlnGlnAsnGlySerValLeuSerGluAspLysSerLysArgLeuAsnThr	100
DB	329	GCCTTCAGCAAAATGGGCTCTTCAGTCTCTCAGAGAACAGCAAGCGTTGAGACACA	388
QY	101	IleLeuAsnThrMetSerThrIleTyrSerThrGlyValLysAsnProAspAsnPro	120
DB	389	ATTCTAAATACATGACACCATCTACAGTACGAGAAAGTTGTAAACCCAGATTAATCCA	448
QY	121	GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140
DB	449	CAGCAATGCTTATCTTACTGAAACAGGTTGAATGAATATATGGCAAAAGATTAGACTAC	508
QY	141	AsnGluArgLeuThrPalatArgLysSerThrPheSerGluValGlyLysGlnLeuArgPro	160
DB	509	AATGACAGGCTCTGGCGCTGGGAAAGCTGAGATCTGAGCTGCGGCAAGCGTGAAGCCA	568
QY	161	LeuTyrGlnGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180
DB	569	TTATATGAGAGATATGCTGTGAAATAATGACATGCGACAGACCAATCTTTATGAGGAC	628
QY	181	TyrGlyAspTyrTyrPheGlyAspTyrGluValAsnGluValAspGlyTyrAspTyrSer	200
DB	629	TATGAGGATATATGAGAGAGACTAGTAAGTAAATGGGTAGATGCTATGACTACAGC	688
QY	201	ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu	220
DB	689	CGCGCCAGTGTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG	748
QY	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro	240
DB	749	CATCTTCACTGCTTATGAGGAGCAAAAGTTGATGATGATGATGATGATGATGATGATG	808
QY	241	IleGlyCysLeuProAlaHisLeuLeuGlyAspMetThrPheArgPheThrAsnLeu	260
DB	809	ATTGAGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	868

QY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
DB	869	TACTCTTGACAGTCTCCCTTTGGACAGAAACCAAACTATATGTTACTGATGCAATGCTG	928
QY	281	AspGlnAlaTyrAspAlaGlnAlaArgIlePheLysGluAlaGluLysPhePheValSerVal	300
DB	929	GACGAGCGCTGGGTGACAGAGATATTCAGAGAGGCGGAGAAATCTTCTTATGCTGTT	968
QY	301	GlyLeuProAsnMetThrGlnGlyPheThrProLysAsnSerMetLeuThrAspProGlyAsn	320
DB	969	GCTCTTCTTATATGATGATCAAGATTTCTGGGAAATTCATGCTAACGAGCAAGAAAT	1048
QY	321	ValGlnLysAlaValCysHisProThrAlaThrAspLeuGlyLysGlyAspPheArgIle	340
DB	1049	GTTCAAGAAAGACGCTGCAATCCACAGCTGGGACCTGGGAGGCGGACCTCAGGATC	1108
QY	341	LeuMetCysThrLysValIleThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis	360
DB	1109	CTTATGTCACAAAGGTGACAAATGACGACTTCTGACGCTCATGATGAGATGGGCAAT	1168
QY	361	IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380
DB	1169	ATCCAGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1228
QY	381	GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu	400
DB	1229	GGATTCATGAAAGCTGTTGGGAAATCATGCTTCTGACGCTCAGCAACCTAACCAATTTA	1288
QY	401	LysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGluThrGluIleAsnPhe	420
DB	1289	AAATCATGTTGGCTTCTCTCTCACCAGTTTCAAGACATGAAACAGAAATTAATCTTC	1348
QY	421	LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys	440
DB	1349	CTGCTCAAAACAAAGCACTCAGATTTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTG	1408
QY	441	ThrArgThrPheValPheLysGlyLysIleProLysAspGlnThrPheLysValTyrTrp	460
DB	1409	TGGAGTGGATGCTTTTAAAGGGAATTTCCCAAGACCGAGTGAGTAAAGTGGTGG	1468
QY	461	GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys	480
DB	1469	GAGATGAGAGCGAGATATGTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1528
QY	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg	500
DB	1529	GACCCCGATCTCTGCTCAGATTTCTATGATTAATCTATTCATTCATTCATTCATTC	1588
QY	501	ThrLeuTyrGlnPheGlnPheGlnLysAlaLeuCysGlnAlaAlaLysHisGluGlyPro	520
DB	1589	ACCCCTTACCATTTCCAGTTTCAAGACACTTGTCAAGCAGCTAATCAATGAAGGCCCT	1648
QY	521	LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu	540
DB	1649	CTGCACAAATGTGACATCTCAAACTCTACAGAAAGTGGACAGAAACTTTCAAAATGCTG	1708
QY	541	ArgLeuGlyLysSerGluProThrPheLeuAlaLeuGluAsnValIleValLysAsn	560
DB	1709	AGCCTTGGAAATTCAGAACCTTGACCTGACCTGAGTGAATGTTGTGAGGACAAAGAC	1768
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln	580
DB	1769	ATGAATGTAAGGCACTGCTCAAGTACTTGAAGCCCTTATTTACTGCTGCTGAAGACAG	1828
QY	581	AsnLysAsnSerPheValGlyTyrPheThrAspTyrPheProTyrAlaAspGlnSerIle	600
DB	1829	AACCAAAATCTTCTTGTGGATGAGTACGCTGAGTCCATTAAGCAGACCAAAAGCATTC	1888
QY	601	LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrPheAsnAsp	620
DB	1889	AAAGTAGAGATTAAGCTTAATAATCAGCTTGTGAGATTAAGCATTAAGATGAGACAGAT	1948

QY 621 GJMETYrLeuPheAspSerSerValAlaTyrAlaMetArgGlnTyrPheLeuVal 640
DB 1949 GAAATGTACTCTTCGATCATCTGTGATATGCTATGAGGCACTACTTTTAAAGTA 2008
QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArg 660
DB 2009 AAAATTCAGATATCTTTTGGGAGGAGATGTGCGATGCGTAATTTGAACACAGA 2068
QY 661 IleserPheAsnPhePheValThrAlaProLysAsnValSerAspIleLeuProArgThr 680
DB 2069 ATCTCTTAATTTCTTTCTGCTACCTACCTAAAGTGTGATATCATCTCTAGAACT 2128
QY 681 GJUALGlnLysAlaIleArgMetSerArgSerArgIleAsnAsnAlaPheArgLeuAsn 700
DB 2129 GAGATGAAAGAGCCATCGAGATGTCCGAGCCGATCAATGATGCTTCCGTCTGAT 2188
QY 701 AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
DB 2189 GACACAGCCTAGAGTTCTGTGGGATACAGCCACCTTGACCTCTTACACAGCCCT 2248
RESULT 8
HSM800880 3341 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA; cDNA DKFZp434A014 (from clone DKFZp434A014);
DEFINITION partial cds.
ACCESSION AL110224
VERSION AL110224.1 GI:5817159
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3341)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submision
JOURNAL Submitted (15-AUG-1999) MIPS, Am Klopferplatz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp434A014) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/
Location/Qualifiers
1..3341
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434A014"
/tissue_type="testis"
/clone_lib="434 (synonym: hhes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2417
/gene="DKFZp434A014"
/gene="DKFZp434A014"
/note="strong similarity to peptidyl-dipeptidase A"
/product="hypothetical protein"
/protein_id="CAB53682.1"
/db_xref="GI:5817160"
/translation="SSSWLLSLVAVTAHSTIEBOATFLDKFNHEADLFYOSL
ASMTNTITTEENOVNNAAGKSAFLKEOSTLAOMPLADEIONLTYVALQALQON
GSSVSEKSKRLNTLNTSTSTSTGKVCNPDPOECLELLEPGLNEIMANSLDYNR
IMAWMSRSEVCKOLRYLXEEVYVLKNEARANHEDYDGRGVYEVNGVDGYYSR
GQIIEVETHEFEIKRYLHIAIVAKDAQIKFEAKFVSVGLPNTQGVWNSMLTD
LYSLVPEQKPNIDVTDAVQADAKIETFAKFEVSVGLPNTQGVWNSMLTD
PGNVKAVCHPTAMDLSKGPRIIMCTKTVMDFLAHHEMHIIDYADVAAPPELLR
NGANEGFEAVGEIMSLSATPKHLKSLGLSPDOEDNETEINFLKQALITVGLP
FTYMLEKRWMMVFKEIIPKDDMMKKMKWEKRELIVGVVEVPHDEYTCDPASLFSVND

YSFTRYRTILXOPFOEALCOAKHEGRLHKRDISNTEBAGOKLENNLRIGKSEPT
LAEENVYGAKKMMVRLPLNTPEPLFTLTKQKNKSRFVNSTDPSPADOSIKVRLSK
SALDRAKAYEMNDNEMYLERSSVAAYOLKVKNNMILFGEEDVARANKPRISENF
EVYAPKVMVSDIIPTEVEKAI RMSRSRININAFRINDNSLEFGLIOPLGPNPPVSI
WLIVFGVKNVGIYVGIYLLFTGIRDRKKRKNKRSRGENPYASIDISKENNPGFQWTD
DYORSF"
polysignal 3266..3271
polya_site 3281
BASE COUNT 1058 a 630 c 740 g 913 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3341
Score: 3869.00 Matches: 720
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x HSM800880 (1-3341)
QY 1 SerThrIleGlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGluAsp 20
DB 54 TCACACATTTAGGAAGAGCCAGACATTTTGGACAAAGTTTACACAGAGCGGAAGAC 113
QY 21 LeuPheTyrGlnSerSerLeuAlaSerThrAsnTyrAsnThrAsnIleThrGluLysAsn 40
DB 114 CTGTTCTATCAAAATTCACCTTCTGTAATTTAAACCAATATTTACTGAGAGAAAT 173
QY 41 ValGlnAsnMetLysAsnAlaGlnLysAspLysThrSerIlePheLeuLysGlnGlnSerThr 60
DB 174 GTCCAAACATGAAATATGCTGGGAGCAAAATGGCTGCTTTTAAAGACAGCTCCACA 233
QY 61 LeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB 224 CTGGCCAAATGATATTCACCTCAAGAAATTCAGATCTCACAGTCAAGCTTCAGCTCAG 293
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysAlaGluAsnThr 100
DB 294 GCTCTTCAGCAAAATGGCTTTCAGTGTCCAGTCCACAAACAGACAAACGCTGCACACA 353
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
DB 354 ATTCATTAATCAATGAGACCATCATCTACGTAAGTAAGTTTAAACCATGATATCA 413
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
DB 414 CAAAGATGCTTATTTACTTGAACCAAGGTTGAATGAAATATGCAAAACAGTTTACTGACTAC 473
QY 141 AsnGlnAlaGluLeuThrPalatTTPGluSerThrArgSerGluValGlyLysGlnLeuArgPro 160
DB 474 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGGGCGGCAAGCTGAGGCCA 533
QY 161 LeuTyrGlnGluTyrValValLeuLysAsnGlnLeuIleAlaArgAlaAsnHisTyrGluAsp 180
DB 534 TTATATGAGAGTATGTGTGCTTGAATAATAGATGCGCAAGCAAAATCATATAGAGAC 593
QY 181 TyrGlnAspTyrTyrPArgLysAspTyrGluValAsnGlnLysValAspGlyTyrAspTyrSer 200
DB 594 TATGGGAGTATTTGGAGAGGAGACTATGAACTAAATGGGAGTAGATGCTATGACTACAGC 653
QY 201 ArgGlnGlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlu 220
DB 654 CGCGGCCAGTTGATGAGAGATGTGAAATCTTTGAAGATTAACCATTTATAGAA 713
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
DB 714 CATCTTATGCTATGATAGTATGAGGCAAGTATGATGATGCTTCTCTCTATATAGTCCA 773
QY 241 IleGlnCysLeuProAlaHisLeuLeuGlnLysAspMetTyrGlyArgPheTyrPheAsnLeu 260
DB 774 ATGGATGCTCCCTGCTCATTTGCTGTGTATGTGGGTAGATTGAGCAAAATCTG 833

QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 834 TACTCTTTGACAGTTCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGGTGT 893
 QY 281 AspGlnAlaTrpAspAlaGlnArgTllePheLysGlnIleValLysPhePheValSerVal 300
 DB 894 GACCGAGCCTGGGATGACACAGAGATATTCAGAGGCGCGAGAAAGTTTGTGATCTGT 953
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 320
 DB 954 GGTCTTCCATATATGACATCAAGAGATTCGGGAAATTCATGCTAACGAGCAGCAAGAAAT 1013
 QY 321 ValGlnLysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTlle 340
 DB 1014 GTTCGAAAGACAGCTGCCATCCACACCTTGGGACCTGGGAGAGCGGCACTTCAGCATC 1073
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHis 360
 DB 1074 CTTATGTGCACAAAGGTGACATGGACACTTCCGACAGCTCATCATGAGATGGGCAT 1133
 QY 361 IleGlnTrpAspMetAlaTrpAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGln 380
 DB 1134 ATCCAGTMTGATGATGCAATGCTGACACACCTTTCTGCTAAGAAATGAGCATATGAA 1193
 QY 381 GlyPheHisGlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1194 GGATTCATGAAAGCTGGTGGGAAATCATGTCTCTGACACACCTTAACCATTTA 1253
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGlnIleAsnPhe 420
 DB 1254 AATATCATTTGGTCTTCTGTCACCGATTTTCAGAGAAACATGAAACGAAATTAACCTC 1313
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys 440
 DB 1314 CTGCTCAAAACAAAGCACTCAGCATTTGTTGGACTCGCCACTTACTTACATGTAGAGAG 1373
 QY 441 TrpArgTrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpTrp 460
 DB 1374 TGGAGGTGGATGGCTTTAAAGGGGAAATCCCAAGACAGTGGATGAAAAAGCTGGTGG 1433
 QY 461 GluMetLysArgGlnIleValGlyValValGlnProValProHisAspGlnThrTyrCys 480
 DB 1434 GAGATGAAGCGAGATAGTGGGGTGGTGGAACTGTGCCCCATGATGAACATTACTGT 1493
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1494 GACCCCGCATCTGTTCCATGTTCTTAATGATTAATCATTCATTCGATTTACACAAG 1553
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuLysGlnAlaIleLysGlnGlyPro 520
 DB 1554 ACCCTTTCACAAATTCACATTTTCAGAAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCT 1613
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlnLysLeuPheAsnMetLeu 540
 DB 1614 CTGCACAAATGTGACATCTCAAACTCTACAGAACTGACAGAAACGTTCATATATCTG 1673
 QY 541 ArgLeuGlyLysSerGlnProTrpThrLeuAlaLeuGlnLysValValGlyAlaLysAsn 560
 DB 1674 AGGCTTGGAAATATGAAACCTGACCTTACATTTGAAATGTTTGAAGGACAAAGAAC 1733
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln 580
 DB 1734 ATGAAATGTAAGCCACAGCTCACTTTCAGCCCTTAATTTTACCTGGCTGAAGACAG 1793
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 DB 1794 AACAAAGATTTCTTTTGGATGAGATGACGATCGGATGCTCATATGACAGCAAGCATC 1853
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysAspLysAlaTyrGlnTyrPasnAspAsn 620
 DB 1854 AAATGAGATTAACGCTTAATATCAGCTTCTGAGATTAAGCATATGAATGAAGACGACAT 1913
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640

DB 1914 GAAATGATCCGTTCGATCATCTGTGCTATGATAGGACGATCTTTAAAGTA 1973
 QY 641 LysAsnGlnMetIleLeuPheGlyGlnLysAspValArgValAlaAsnLeuLysProArg 660
 DB 1974 AAAATATGATGATCTTTTGGGAGAGATGCTCGATGCTTAATTTGAACCAAA 2033
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2034 ATCTCCCTTAATTTCTTTGTGCTGCTGACCCGTAATATGCTGATATCTTCTAGACT 2093
 QY 681 GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2094 GAAATGAAAGCCATCAGAGATGCTCCGAGCCGATGCAATGATCTTCCGCTGAAAT 2153
 QY 701 AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2154 GACAAAGCCTTAGATTTCTGGGATACAGCCAAACCTTGACCTTCAACAGGCCCT 2213
 RESULT 9
 ARI35177
 LOCUS ARI35177 3396 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 1 from patent US 6194556.
 ACCESSION ARI35177
 VERSION ARI35177.1 GI:14124082
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3396)
 AUTHORS Action, S. Laurene, and Robison, K. Earl.
 TITLE Angiotensin converting enzyme homolog and therapeutic and diagnostic uses therefor
 JOURNAL Patent: US 6194556-A 1 27-FEB-2001;
 FEATURES location/Qualifiers
 source 1..3396
 BASE COUNT 1034 a 659 c 772 g 931 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x ARI35177 (1-3396)
 QY 1 SerThrIleGlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAsp 20
 DB 136 TCACATTTGAGGAAACAGGCCAAGACATTTTGGACAGTATTAACACAGAGCCGAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGlnLysAsn 40
 DB 196 CTGTCTATCAAAAGTTCACCTGCTTCTGGAATTAACACCAATATTAAGTGAAGAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnIleSerThr 60
 DB 256 GTCCAAACATGAAATATCTGCGGCAAAATGCTGCTTTTAAAGAACAGTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 316 CTGGCCCAATATGATCATCTACAGAAATTCGAATTCACAGTCAAGCTTCAGCTCAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100
 DB 376 GCCTTCACCAAAATGGCTTCAGTGTGCTCAGAAAGACAAAGACGTTGAACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 436 ATTTCAATATCAATGAGCAGCATCTACAGTACTGGAAGAAAGTTTGAACCCAGATATCA 495

QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluLeuAlaAsnSerLeuAspTyr 140
 DB 496 CAGAAAGCTTACTTACTGAAACAGGTTGAATGAATAATGCAAAACAGTTAGACTAC 555
 QY 141 AsnGlnArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyGlnLeuAlaPro 160
 DB 556 AATGACAGGCTTGCGCTTGGAAGCTGGAGATGTGAGGTGGCAAGCAGCGAGAGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTAATGAAGAGATATGTCCTTGAAATAATGAGATGGCAAGACCAATCATTTATGAGGAC 675
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 676 TATGGGATTTATGGAGAGAGACTATGAGTAAGTAAGGGGTAGATGGCTATGACTACAGC 735
 QY 201 ArgGlyGlnLeuLeuGluAspValGluHisThrPheGluGluLeuLysProLeuTyrGlu 220
 DB 736 CCGGCGCAGTTGATTAAGATGTGGAACATACCTTTGAAAGAGATTAAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCCATATGTGAGGCGCAAGTTGATGAATGCCATACCTTCATATACATCCA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTrpHisLeu 260
 DB 856 ATTTGATGGCTCCCTGCTCATTTGCTGGTGTGATGTGGGTAGATTGGCAAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValHisAlaMetVal 280
 DB 916 TACTCTTGACAGTCCCTTTGACAGCAACCAACATATGATTTACTATGCAATGAGTG 975
 QY 281 AsnGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPheMetValSerVal 300
 DB 976 GACACAGGCTGGGATGCCAGAGAAATATTCAGAGAGCGGAGAAAGTCTTTGATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGGCTTCCTAATATGCTCAGAGATTTCTGGAAATTCATGCTAAACGAGCCAGAGAAAT 1095
 QY 321 ValGlnLysAlaValCysHisSerThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1096 GTTCAGAAAGCAGTGTGCATGCCACAGCTGGGAGCTGGGAGGCGCACTTCAGAGATC 1155
 QY 341 LeuMetCysThrLysValIleMetLysPheLeuThrAlaHisHisGluMetGlyHis 360
 DB 1156 CTATATGTGCAGAAAGGTGCAATGACGACTTCCTGACAGCTCATCATGAGATGGGCAAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1216 ATCCAGTATGATATGCAATATGCTGTGCACACCTTTTCTGCTAAGAAATGGAGCTATGAA 1275
 QY 381 GlyPheHisGluAlaValGlyIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCACATGAAGCTGTGGGAAATCATGTCATCTTCGACACGCCACACCTAAGATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AATTCATATGGTCTTGTGTGCACCCGATTTTCAAGAGACAAATGAAGAAATTAATCAATTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAAAACAGCACTCAGCATTTTGGGACTGTGCCATTTACTTACATGTTATGAGAG 1455
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp 460
 DB 1456 TGGAGGTGAGATGCTTTAAAGGGGAAATTCACAAAGACAGATGATGAAAGAGGTGG 1515
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAAGCCGAGATATGTTGGGTGGAAACCTGTGCCCATGATGAAGAAACATATACGT 1575

QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1576 GACCCCGCATCTCTGTTCATCTTTCTAATGATTTACTCATTTTCATTCATTCATTCACAAAG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCAATTCACATTCAGTTTCAGAGACACTTTTGTCAAGACGATCAAAACATGAAGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTTCAATATGTGTG 1755
 QY 541 ArgLeuGlyLysSerGlnProTrpPheThrLeuAlaLeuGluAsnValValGlyValLysAsn 560
 DB 1756 AGCCTTGGAATTCAGAACCCCTGGACCTAGCATTTGGAAATGTTGTGAGAGCAAGAAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTyrPheLysAspGln 580
 DB 1816 ATGAATGTAAAGGCCACTGCTCAACTACTGAGCCCTTATTTACCTGGCTGAAAGACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrArgPheTrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAAAGATTTCTTTGTGGATGGAGTACCGACTGGAGTCCATATGACAGCAACCAAGATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
 DB 1936 AAAGTGAGATTAAGCTTAATATCAGCTTGTGAGATTAAGCATATGATGAGACAGACAAAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACCTGTTCCGATCATCTGTTCCATATGATGTATGAGGAGCACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAATATCAGATGATCTTTTGGGAGAGAGATGTCCAGAGTGGCTAATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTAAATTTCTTTGTCTGACCTGACCTTAATATGATGTGATATCATTTCCATGAACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAGCTTGAAAGGCCATCAGATGATGCCGAGCCGTATCAATGATGCTTCCGCTCAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACAACACCTCAGAGTTCTTGCGGATACAGCAACACTTGAGACTCCTTAACAGACCCCT 2295
 REFSEQ 10
 LOCUS AX418982 3396 bp DNA linear PAT 18-JUN-2002
 DEFINITION Sequence 1 from Patent WO0212471.
 ACCESSION AX418982
 VERSION AX418982.1 GI:21523781
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 ACTION, S., Robison, K.E. and Hsieh, F.Y.
 Angiotensin converting enzyme homolog and uses therefor
 Patent: WO 0212471-A 1 14-FEB-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..3396
 location/Qualifiers
 1..3396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 82..2499
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD35458.1"
 /db_xref="GI:21523782"

```

/translation="MSSSMLLSLVAVTAOSTIEQAKFIDKFNHAEPLFYOSS
LASMNNTNTEENONNMNAGDKMSAFIKKQSTLQAOPLQIOLQLOALOO
KSSSYSEDKSKRNTIITNTASTYSTGVKCPNDPOCLLEPGINTEANSLDNE
KRWAMSESRSEKOLRPLEETVALKNEMANVNTEDGYWRDYEYNGVDGYS
KQLJEDVEHTEEEKPLXELHAYRAKLNAPSYISPLCDPAHLHGDGMREWT
NLSYLVPGOKPNDIDVANDVQANDORIFKEAKEFVSGLNMTGDEMENMLT
DGNVOKAVCHPTAMDLDGDFRILMCTVMDLFAHNEHGHIOYMAVAOQPLL
RNGANGPEHAYGEIWSLSAATPKHLKISGLSPDEBNETEINELKQALTYGTL
PRYMLERKRNKAVFGEIPIKDOMKMKMKREIYGVPEPPHDETCDDPASTFHSN
DYSEITRYTRTXKQFOPALCOAKHBEPLKICDINSSTEGAKTEPNLRIGKSEPW
TALLENVGAKNMNAEPLNTEPLEFTWLKDNKSEFMSWSTDMSPADQSIAKVISL
KSAIDKAYEMNDENMYLFRSSVAYAMROYFLKVNQMILLFEQVPRANKPRISFN
FEVTAKNVSDIIPTEVEKAIMSRSRINDAFRLNDSLEIAGILOPTGPNOPVS
IMLVGVAVGIVGIVILIFITGIRDKKKKNSGENPYASIDISKENPNPQONT
DVOQTSF"
BASE COUNT      1034 a      659 c      772 g      931 t
ORIGIN

Alignment Scores:
Pred. No.:      0      Length:      3396
Score:          3869.00      Matches:      720
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:               6      Gaps:      0

US-09-978-385-2_COPY_19_738 (1-720) x AX418982 (1-3396)

QY      1 SerThrIleGluLuglnAlaLysThrPheLeuAspGlyPheAsnIleGluIleAsp 20
DB      136 TCCACCACTTGAGAAAGCCCAAGACATTTTGGACAAGTTTACACGACGACCGCAAGAC 195
QY      21 LeuPheTyrGlnSerSerLeuAlaSerTyrPAsnTyrAsnThrAsnIleThrGluIleAsn 40
DB      196 CTGTCTATCAAAAGTCTACTGCTCTTCTGGAATTAACACCAATATTAATCTGAAGAAGAT 255
QY      41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuGluGlnIleSerThr 60
DB      256 GTCCAAACATGATATAATGCTGGGGACAAATGGCTGCTTTTAAAGGAACGTCACCA 315
QY      61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnIleGln 80
DB      316 CTTCGCCAAATGATCCACTACACAAAGAAATTCAGATCTCAGTCAAGCTTCAGCTGACG 375
QY      81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
DB      376 GCTCTCAGCAAAATGGCTCTTCACTGCTCTCAGAACACAAAGCAACGTTGAACACA 435
QY      101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
DB      436 ATTCTAAATACAAATGACACCACTCTACAGTACTGGAAGAAATTTGTAAACCCAGATATCCA 495
QY      121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
DB      496 CAACAATAGCTTATTAATCTGAACCAAGTTTGAATGAATTAATGGCAACAGTTTAACTAC 555
QY      141 AsnGluArgLeuTrrAlaTrrPglUserTrrArgSerLysValGlyLysGlnLeuAspPro 160
DB      556 AATAGAGGCTCTGCTGGGAAAGCTGAGATCTGAGGTGGCAAGCAGCTGAGGCCA 615
QY      161 LeuTyrGlnGluTyrValIleValLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
DB      616 TTATATATGAGAGTATGGTCTTGAATAATGACAGTGCAGAGAAATCATTTATGAGGAC 675
QY      181 TyrGlyAspTyrTrrArgGlyAspTyrGlyValAsnGlyValAspGlyTyrAspTyrSer 200
DB      676 TATGGGAGTATTTGGAGAGGACTATGAGATTAATGGGGATGAGTATGACTACAGC 735
QY      201 ArgGlyLysLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGlu 220
DB      736 CGCGGCGCAGTGTGATGGAAGATGTGAACATACCTTTGAAGAGATTAACCATTAATAGAA 795
QY      221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240

```

Db 1876 AACAGAAATCTTTTGGGATGAGTACCGACGATGATCCATATGACACCAAGCATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAsn 620
 Db 1936 AAAGTGAAGATTAACCCCTTAATGCTCTGGAGATTAACCATATGATGAAAGCAAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuVal 640
 Db 1996 GAAATGATCCTGTTCCATCATCTGTCGATATGCTATGAGGAGCTACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValAlaTyrAlaAsnLeuLysProArg 660
 Db 2056 AAAAATACATGATCTCTTTTGGGAGAGAGATGCGAGTGGCTAATTTGAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2116 ATCTCTTAAATTTCTTGTCTGCTGACCTGACCTAAAGTGTCTATATCATCTTAAAGT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2176 GAAGTGAAGAGCCATCAGATGTCGCCGAGCCGATCATGATGCTTCCGCTGAAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACACAGCCTGATGTTCTGGGAGATACAGCCACACTTGGACCTCTTAACGACCCCT 2295
 RESULT 11
 AX431513
 LOCUS AX431513 3396 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 1 from Patent WO0239997.
 ACCESSION AX431513
 VERSION AX431513.1 GI:21656345
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Gould, A.E., Stricker-Krongrad, A., Acton, S.L., Brown, J.A., Guan, B.,
 Dales, N.A., Kadambi, V.J., Ocalin, T.D., Patane, M., and Solomon, M.,
 Ace-2 modulating compounds and methods of use thereof
 JOURNAL Patent: WO 0239997-A 1 23-MAY-2002;
 MILLENNIUM PHARM INC (US)
 FEATURES
 source Location/Qualifiers
 1..3396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 82..2499
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37288.1"
 /db_xref="GI:21656345"
 translation="MSSSSKLLSLVAVTAOSTTEQAKTFLDKFNEHAEDELYQSS
 LASMTNTNTEENVQNNMAGDKMSAFLEKQSTLAQYPLQEQNLTKLQALQO
 NGSSVLSBDSKRLNTLNTSTIYSGKCNPNPOCLLEGLNEIMANSIDYNE
 RLMAWSESRSEVQKRLPLEYELVYLKNEMARHNVDEYGYGVNDVVDYDS
 RGQLEDEVEHTFEELKPLEYELHAYVAKLNNAYPSYSGICPAHLGKMRMT
 NLYSLTVEFGQKPNIDYDAMVQDAQDAORLEKAEKFTVSVCIPNMTQGWENSMIT
 DPGNVOKAVCHPTAMDAGKGFRLMCTKVTMDPLRAHHEMGHIQYDMAVAPFL
 RNCANEGHEHVEGELMSLAATPKHLKSLISGLSPQEDNTEINFLKQALITVGL
 PFTTLEKMRMVEKGETPKQDMKRWEMKREIVGVEVPHDETCDPASTLPHVSN
 DYSLTRYTTRLVQFOEALQAKHGGPLHKCDISNTEAGOKLFNMLGKSEYM
 TLAENVVGAKNMVRPLNTEFELFTLKQNKNSPYGMSDMSPYADDSIKRIST
 KSALEGDKAYENWENDEMLERSSVAYAMROYLKYKNOMIILEGDEVDVANKPRISN
 FEVTAPKRVSDIIRTEVEKAKIRMSRINDAPFLNDSLEFLIQFLPGPNPPVS
 IMLTIVGVKVIYVGIIVLITGIRKRRKKNKNSENPNASIDISKGENNPGFQNT
 DDVQTSF"
 BASE COUNT 1034 a 659 c 772 g 931 t
 ORIGIN
 Alignment Scores:
 Score: 0 Length: 3396
 3869.00 Matches: 720

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AX431513 (1-3396)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHISgluAlaLysP 20
 Db 136 TCCACCATGAGAGAACAGCCAAAGACATTTTGGACAAGTTTAAACAGCAAGCCAAAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerThrPheAsnTyrAsnIleThrGluLysAsn 40
 Db 196 CTGTTCTTACAAAGTTCACCTCTCTGGAATTTAAACCAATTTATCAAGAGCAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrPheAlaPheLeuGlnGluSerThr 60
 Db 256 GTCCAAAACATGAATAATGCTGGGACAAATGCTGCTTTTAAAGCAACAGTCACCA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 316 CTTCGCCAAATATATCACTACACACAAATTCAGATCTCACAGTCAAGCTTCAGCTGAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 376 GCTCTTCAGCAAAATGGGCTCTTCAGTCTGTCAAGACCAAGCAACGGTTGACACCA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 436 ATTCATTAATACAAATGACCAACCATCATCAAGTCTGGAAGAAAGTTTAAACCCAGATTAACCA 495
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 Db 496 CAAGAACTCTTATTAATCTTGAACCAAGGTTGATTAATTAAGCAACAGTTTATGACTAC 555
 QY 141 AsnGluArgLeuThrPalaTrpGluSerTyrArgSerGluValIleGlyLysGlnLeuAspPro 160
 Db 556 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTGGCAAGCAAGCTGAGGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHISArgLysP 180
 Db 616 TTATATGAAGAGTATGAGTCTTGAAGAAAGAGATGCGCAAGCAAGCAATCATTTATGAGAC 675
 QY 181 TyrGluAspTyrTyrPalaGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 676 TATGGGATTTATTTGAGAGAGGAGACTATGAATGAATGGGATGAGTGTGACTATGACG 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGlnIleLysProLeuTyrGlu 220
 Db 736 CGCGGCAAGTTGATGAAGATGTGGAACATACCTTTGAAGAGATTAACCATTAATGAA 795
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAsnAlaLysProSerTyrIleSerPro 240
 Db 796 CATTTTATGCTTATGATGAGAGGCAAAAGTATGATGATGCTTCTTCATATCACTTCA 855
 QY 241 IleGlyCysLeuProAlaHisIleLeuGlyAspMetTrpLysArgPheTrpThrAsnLeu 260
 Db 856 ATTGAGATGCTCTCCCTGCTATTTGCTGGATATGATGGGGTATGATTTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValAlaProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 916 TACCTTTCAGAGTCTCTTTCGACAGCAAAACCAAAATAGATGTTACTATGCAAGTGG 975
 QY 281 AspGlnAlaTrpAspAlaGlnAlaGlyIlePheLysGlnAlaGluLysPhePheValSerVal 300
 Db 976 GACACAGCCGCGGATGACAGACAAATATTAAGAGAGCCGACAGATTTCTTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 1036 GGCTCTTCTTAATATGACTCAAGGATTTCTGGAAATTTCACTATGACGACCCAGCAAT 1095
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyLysPheArgTle 340

Db 1096 GTTCGAAGACGCTGCCATCCACAGCTTGGGACCTGGGGAAGGGGAGCTTCAGATC 1155
 Oy LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisGluMetGlyHis 360
 Db 1156 CTTATGTGCACAAAGATGACATGACACTTCCGACAGCTCATCAAGATGGGGAT 1215
 Oy 361 IleglnIyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuAsnGlyAlaAsnGlu 380
 Db 1216 ATCCAGATGATGATGGCATATGCTGCACAACTTTTCTGCTTAAGAAATGGAGCTATATA 1275
 Oy 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1276 GGATTCAGATGAGCTGGTGGGAATCATGTCATCTTCTGAGCCACACTTAAGATTTTA 1335
 Oy 401 LysSerIleGlyLeuLeuSerProAspPheGlnIleuAspAsnGluThrGlnIleAsnHe 420
 Db 1336 AAATCCATGGCTCTTCTGTCACCCGATTTTCAAGAAAGCAAGAAACGAAATTAACCTTC 1395
 Oy 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 Db 1396 CTGCTCAAAACAGACACTGATGTTGGGACTCTGCCATTTACTTACTGTTAGAGAAG 1455
 Oy 441 TrrArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrp 460
 Db 1456 TGGAGGTGATGCTTTTAAAGGGGAAATCCCAAGACAGCTGATGAAAGAGTGGTG 1515
 Oy 461 GluMetLysArgGluIleValGlyValGlyProValProHisAspGluThrTyrCys 480
 Db 1516 GAGATACAGCGAGATGATTTGGGGTGGGAGAACCTGCTGCCCATGATGAACATACCTCT 1575
 Oy 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 Db 1576 GACCCGCACTCTGCTTCATGTTCTTAATGATTCATTCATTCATTCATTCATTCATTC 1635
 Oy 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaHisGluGlyPro 520
 Db 1636 ACCCTTACCAATTCAGATTCAGAAAGCACTTTGCAAGCAGCTAAACATGAGGCCCT 1695
 Oy 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540
 Db 1696 CTGCACAAATGATGACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTTCAATATGCTG 1755
 Oy 541 ArgLeuGlyLysSerGluProTrpPheThrAlaLeuGluAsnValGlyAlaLysAsn 560
 Db 1756 AGCCTTGGAAATCAAGAACCTGACCTGACCTGAGCTTGGAAATGTTTGGAGCAAGAAC 1815
 Oy 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
 Db 1816 ATGAATGTAAAGCCACTGCTCACTACTTGAACCCCTTATTTACTGCTGGAAGACACAG 1875
 Oy 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 Db 1876 AACCAAGAAATCTTTTGGGATGAGTACCGACTGGAGTCCATATGCAAGCAACCAACATC 1935
 Oy 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
 Db 1936 AAGGTGGATTAAGCTTAATCAAGCTCTGGAGATTAACCAATGATGATGAGCAAGCAAT 1995
 Oy 621 GluMetLysLeuPheArgSerSerValAlaTyrAlaMetArgLysTrpPheLeuLysVal 640
 Db 1996 GAAATGTACCTGCTCCGATCATCTGTCATATGCTATGAGGACAGTACTTTTAAAGTGA 2055
 Oy 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 Db 2056 AAAATATCAGATGATCTTTTGGGAGAGAGATGTCGATGCTTAATTTGAAACCAAGA 2115
 Oy 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2116 ATCTCCTTAATTTCTTGTGCACCTGACCTAAATATGCTGATATCATTTCTAGAACT 2175
 Oy 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2176 GAAGTTGAAAGGCCATCAGATGTCCCGAGCCGATATCAATGATGCTTCCCTGTAAT 2235

Oy 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACACAGCCTATAGATTTCTGGGGATACAGCCAACTTGAGACTCTTAACAGACCCCT 2295

RESULT 12

AF241254

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL</

QY 1 SerThrIleGluGluGluGluAlaIysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 158 TCCACCATTGAGGAAGAGGCCAAGACATTTTGGACAGATTAAACACAGAAAGCCGAGAC 217
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnHisIleThrGluGluAsn 40
 Db 218 CTGTCTATCAAAAGTTTCACTTCTCTTGGATTATACACCAATATTTACTGAGAGAAAT 277
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 Db 278 GTCCAAACATGATATATGCTGGGAGCAAAATGGCTCTTTTAAAGAACAGCTCCACA 337
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 338 CTGGCCCAATGTTCCACTACAGAAATTCAGAAATCTCAGCAAGCTTCAGCTGACAG 397
 QY 81 AlaLeuGlnLysAsnGlySerSerValLeuSerGlnAspLysSerLysAspGluAsnThr 100
 Db 398 GCTCTTCAGCAAAATGGGCTTTCAGTCTCTCAGAACACAGAACAGCAAGCGTTGAAACA 457
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 458 ATTCTAAATCAATGAGCACCATCTACAGTAAGTGAAGAAAGTTGTAAACCAAGATATCCA 517
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetLysAsnSerLeuAspTyr 140
 Db 518 CAAAGATGCTTATTAATCTGACCAAGGTTGAAATGAATGAATGAAGCAAGTTTACACTAC 577
 QY 141 AsnGluArgLeuTyrAlaTyrGlnSerTyrArgSerGlnValGlyGlnLeuArgPro 160
 Db 578 AATGAGAGGCTCTGGGCTGGGAAAGCTGAGATCTGAGTCCGCAAGCAAGCTGAGGCCA 637
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 638 TTATATGAAAGATATGCTGCTGAAAGATGAGATGAGCAAGCAAAATCATTAATGAGGAC 697
 QY 181 TyrGlyAspTyrTyrArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyr 200
 Db 698 TATGGGATATATGAGAGAGAGACATGAAATGAATGAATGAATGAATGAATGAATGAAT 757
 QY 201 ArgGlyGlnLeuIleLysAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 Db 758 CGGGCCAGCTGATGTAACATGAGCAACATGACCTTGAAGAGATTAACCATTAATATGAA 817
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 Db 818 CATCTTCATGCGCTATGAGGCAAGGATGATGATGCTTATCTCATATATCAAGTCCA 877
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrPheAsnLeu 260
 Db 878 ATTTGATGCTCTCCCTGCTCATTTGCTGTGATATGAGTGGGATTTTGGCAAAATCTG 937
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 938 TACTCTTTACAGCTTCCCTTGGACCAAAACCAATAGATGATACATGATCAATATGCTG 997
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 Db 998 GACCGAGGCTGGGATGACAGAGAAATATCAAGAGAGGCGCAAAATTTCTTTGATCTGTT 1057
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPheLysAsnSerMetLeuThrAspProGlyAsn 320
 Db 1058 GGTCTCTCTATATATATCAATCAAGATTTGGGAAATTTCCATGCTTAACGAGCCAGAAAT 1117
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPhePheLysGlyLysGlyAspPheArgIle 340
 Db 1118 GTTCAGAAAGAGCTGTCATCCACACCTTGGGAGCTGGGAGAGGCGAGCTTCAGGATTC 1177
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 360
 Db 1178 CTTAAGTGCACAAAGATGACATGAGCACTTCTGACAGCTCATCATGAGATGGGAGAT 1237
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380

Db 1238 ATCCAGTATGATATGCGCATATGCTGCACAAACCTTTTCTCTAAGAAATGAGACTATGAA 1297
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1298 GGATTCATCAAGAGCTGTGGGGAATCATGCTACTTTCGACAGCCACACTTAACATTTA 1357
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 Db 1358 AATTCATTTGGCTTCTGTCACCCGATTTTCAAGAAAGCAATGAAGAAACAGAAATTAACCTC 1417
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuLys 440
 Db 1418 CTGCTCAACAGACATCACTACGATGTGGGACTCTGCCATTTACTTACATGTTAGAGAG 1477
 QY 441 TyrArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysTyrTyrP 460
 Db 1478 TGGAGGTGGATGCTTTTAAAGGGAAATTTCCAAAGAACCATGATGAAAAAGTGTGG 1537
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
 Db 1538 GAGATGAGCGAGAGATGAGTGGGGTGGGAACTGTGCCCATGATGAAACATATCTGT 1597
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrArg 500
 Db 1598 GACCCCGCATCTCTGTCCATGTTCTATATGATTAATCTCATTCATTCATATACACAGG 1657
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 Db 1658 ACCCTTTTACCATTCCAGTTTCMAAGACACCTTGTCAAGCAGCTGAACATGAAGGCCCT 1717
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyLysLeuPheAspMetLeu 540
 Db 1718 CTGCACAATGTGATCATCTCAAACTCTACAGAGCTGACAGCAAACTGTTAAATATGCTG 1777
 QY 541 ArgLeuGlyLysSerGluProTyrPheLeuAlaLeuGluLysValValGlyAlaLysAsn 560
 Db 1778 AAGCTTGGAAATCAAGAACCTCGACCTTACATTTGAAATGTGTGAGAGCAAGAAC 1837
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGln 580
 Db 1838 ATGATATGAAGCCACTCTCACTTATGAGCCCTTATTTACTGCGTGAAGACAG 1897
 QY 581 AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 Db 1898 AACAGAAATTTCTTTTGGGATGAGTACGACAGTGCAGATGCATATGCAGACCAAGCATC 1957
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysLysAlaTyrGluTyrPheAspAsn 620
 Db 1958 AAGTGAAGATTAACCTTAATATCAAGCTTGGAGATTAAGCATATGAATGGAACGACAT 2017
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 Db 2018 GAAATGATACCTGCTCCGATCATCTGTTCATATCTATGAGCGATCTTTTAAAGTA 2077
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 Db 2078 AAAATATAGATATCTTTTGGGAGAGAGATGTGCAATGGCTAATTTGAAACAGACA 2137
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2138 ATCTCTTTAATTTCTTTGTCACCTGCACCTTAATAATGTGTGATATCTCTTAAGACT 2197
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2198 GAACCTTAAAAAGGCATACAGATGTCCCGAGCGGTATCATATGATGCTTCCTGCTGAAT 2257
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuLysProProAsnGlnProPro 720
 Db 2258 GACACAGCCTTAGATTTCTGGGATACAGCAACACTTGACCTCTTAACAGCCCCCT 2317
 RESULT 13
 AB046569

QY 1 SerThrIleGluGluAlaLysThrPheLeuAspLysPheasnHisGluValAsp 20
 |||||
DB 109 TCCACCACTTGAGGACAAGGCCAAGACATTTTGGACAAGATTAAACCACGAAGCGGAGAC 160

OY	21	LeuPheHyrcInSerSerLeuAlaSerTyrAspTyrAsnThrAsnIleThrGluGluAsn	40
Db	169	CTGTTCATCAAAAGTCTACTTGGCTCTTGGCAATTATACACCAATATTACTGAAGCAAT	228
OY	41	ValGInAsnMetAsnAsnAlaGlyAspLysTyrPSerAlaPheLeuLysGluInSerThr	60
Db	229	GTCCAAAACATGATTAACGCTGGGGGACAAATGGCTCTCTTTTAAAGGAACGTCACAA	288
OY	61	LeuAlaGInMetTyrProLeuGInGluIleGlnAsnLeuThrValLysLeuGInLeuGln	80
Db	289	CTTCCCAAAATGTAATCCACTACAGAAATTCAGATCTCACAGTCAGCTTCACGTCAG	348
OY	81	AlaLeuGInGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr	100
Db	349	GCTCTTCAGCAAAATGGGTCTTCAGTGCCTCCACAAAGACAAAGCAAGCTGGTGAACA	408
OY	101	IleLeuAsnThrMetSerSerThrIleTyrSerThrGlyLysValGlyAsnProAspAsnPro	120
Db	409	ATTCTAAATACAAAGACACCATTCATACAGTCTGAAAGGTTTGTAAACCCAGATAATCCA	468
OY	121	GInGluCysLeuLeuLeuGluProGluLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140
Db	469	CAAGAAATGCTATTACTTGAACCAAGTTTGAATGAATAATGGCAAAACAGTTTAGCTAC	528
OY	141	AsnGluArgLeuThrAlaTyrPglUserTyrArgSerGluValGlyLysGlnLeuArgPro	160
Db	529	AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGACTGGAGTGGGCAAGCAGCTGAGGCCA	588
OY	161	LeuTyrGluGluTyrValValLeuLysAsnAluMetAlaArgAlaAsnHisTyrGluAsp	180
Db	589	TTATATGAAAGATGTGTGCTTGTGAATAATGAAATGGCAAGACCAATCATTAAGAGAC	648
OY	181	TyrGluAspTyrTyrPArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer	200
Db	649	TATGGGAGTATTTGGAGAGGAGACTATGAACTAATGGGGTAGATGGCTATGACTACAGC	708
OY	201	ArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGlu	220
Db	709	CGGGGCCAGTTGATGGAAGATGTGGACATACCTTTGAAGGATTTAAACCTTATATGCA	768
OY	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro	240
Db	769	CATCTTCATGCGCTATGTAGAGGCAAAAGTTGATGTAATGCCCTATCTCTAATACAGTCCA	828
OY	241	IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrPglTyrPheThrThrAsnLeu	260
Db	829	ATTGGATCCCTCCCTGCTCATTTGGTGTGATATGTGGGTAGATTTTGGACAAATGTG	888
OY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	889	TACTCTTTGACAGTCCCTTGGACAGAAACCAATAGATGTACTAGATGCATGTG	948
OY	281	AspGlnAlaTyrPaspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal	300
Db	949	GACCCAGGCGCTGGAGGACAGAGAAATATTCAAAGGGGCCGAGAAATTTTGTGTATCTGT	1008
OY	301	GlyLeuProAsnMetThrGlnGlyPheThrPglAsnSerMetLeuThrAspProGlyAsn	320
Db	1009	GGTCTCTCTTAATATACTCACAGATTTCTGGAAATTTCCATGCTCTAACGACCCAGAAAT	1068
OY	321	ValGlnLysAlaValCysHisProThrAlaTyrPaspLeuGlyLysGlyAspPheArgIle	340
Db	1069	GTTCAGAAAGCAGTCTGCCATCCACACGCTGGGACCTGGGGAAAGGCGACCTTCAGGATC	1128
OY	341	LeuMetCysThrLysValThrMetAspPaspPheLeuThrAlaHisIleGlnMetCysLys	360
Db	1129	CTTAATGTGCAGAAAGGTGACAAATGGACAGACTTCTTGACAGCTCATCAGATGGGGCAT	1188
OY	361	IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380
Db	1189	ATTCAAGTATGATATGGCAATATGCTGCACAACTTTTGTGCTTAAAGAAAGGAGCTAATGAA	1248
OY	381	GlyPheHisGluAlaValGlyLysIleLeuMetSerLeuSerAlaAlaThrProLysHisLeu	400

Db 1249 GGAATTCATGAGCTTGTGGGGAATCATGTCATCTTTCGACGCCACACCTAAGCATTTA 1308
 QY 401 LysSerIleGlyLeuSerProAspPheGlnGluAspGlnThrGluLeuSnphe 420
 Db 1309 AATCATGATGCTTCTTGTCCACCGCATTTTCAAGAGAACAAAGAAACAAATTAACCTTC 1368
 QY 421 LeuLeuGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 Db 1369 CTGCCTCAAAAGAGCACTCAGATGTTGGAGCTCTGCCATTTACTTACTTACATGTTAGAGAG 1428
 QY 441 TrpArgTrpMetValPheGlyGlyIleProGlyAspGlnTrpMetLysLysTrp 460
 Db 1429 TGGAGGTGATGCTCTTTTAAAGGGAAATCCCAAGACCATGGATGAAAAAGTGTGG 1488
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGlnThrTyrCys 480
 Db 1489 GAGATGAAGCGAGAGATGTTGGGTGGGAGAACCTGTGCCCATGATGAACATACATCTGT 1548
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 Db 1549 GACCCCGGACCTCTCTTCATGTTTCTAATGATTCATTCATTCATTCATTCATTCATTCAT 1608
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaValHisGlnGlyPro 520
 Db 1609 ACCCTTTACCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1668
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 Db 1669 CTGCACAAATGTCATCTCAAACTCTACAGAGCTGACAGAACTCTTCAATATGCTG 1728
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAlaValValAlaLysAsn 560
 Db 1729 AGCGTTGAAATATCGAACCTGACCTTACATTCGAAATGTTGAGGACCAAGAAC 1788
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
 Db 1789 ATGATGTAGAGGCACTCTCACTACTTTCAGCCCTTATTTACTTGGCTTCAAAAGACAG 1848
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 Db 1849 AACAGAAATTTCTTTGGGAGAGTAGACGATGAGATGATGATGAGACCAAGATC 1908
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysAspValAlaTyrGluTrpAsnAsp 620
 Db 1909 AAATGAGAGATTAACCTTAATATCAAGCTTGGAGATGAGATGAGATGAGATGAGATGAG 1968
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 Db 1969 GAAATGTACCTGTTCCATCATCTGTTGCATATGCTATGAGGACGATCTTTTAAAGTA 2028
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 Db 2029 AAAATATCATGATTTCTTTTGGGAGAGAGATGTGAGATGCTTAATTTGAAACCAAGA 2088
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 Db 2089 ATCTCTTATATTTCTTTGTCACCTGACCTAAATATGTCATGATTCATTCATTCATTCAT 2148
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 Db 2149 GAAATGTAAAGGCAAGAGATGTCCCGGAGCGCATTCATGATGCTTCCGCTGCAAT 2208
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2209 GACAACAGCTTAGAGTTTCTGGGGATACAGCCAACTTGGACCTCTTAACAGCCCTT 2268
 RESULT 14
 AX463938
 LOCUS AX463938 3732 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 71 from Patent WO0140466.
 ACCESSION AX463938
 VERSION AX463938.1 GI:21898986

KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E.,
 Gao, W.O., Gerritsen, M.B., Goddard, A., Godowski, P.J., Gurney, A.L.,
 Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,
 Wood, W.L., and Zhang, Z.
 TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
 same
 JOURNAL Patent: WO 0140466-A 71 07-JUN-2001;
 Genentech Inc. (US)
 FEATURES
 SOURCE Location/Qualifiers
 1..3732
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1137 a 722 c 821 g 1052 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,296-312 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 6 Gaps: 1
 US-09-978-385-2_COPY_19_738 (1-720) x AX463938 (1-3732)
 QY 1 SerThrIleGluGlnGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 94 TCCACCTTATGAGAGACAGCCCAAGACATTTTGGACAAGTTTAACACAGAACCCGAAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluGluAsn 40
 Db 154 CTGTTCTATCAAAAGTTCACCTGCTCTTGGAAATTATTAACCAACCAATTTCTGAAAGGAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 Db 214 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 274 CTGCCCAATGATATCCATACACAAAGAAATTCAGAAATCTCACAGTCAAGCTTCAAGTCCAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLysSerGluAspLysSerLysArgLeuAsnThr 100
 Db 334 GCTCTTACGCAAAATGGGCTTCAGTCTCTCAGAAAGACAAAGCAACGGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnPro 120
 Db 394 ATCTCAATATACATGATGACACCATCTACATCTGCAAAAGTTGTGAACCCAGATATATCCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 Db 454 CAAGAAATGCTTATTTACTTGAACACAGGTTGAATGAATATGCAAAAGCTTTAATACATAC 513
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 Db 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAATCTGAGTCTGGCAACAGCTGAGGCCA 573
 QY 161 LeuTyrGluGluLysTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 574 TTAATGAAGAGATGATGTGCTTGAATAATGAGATGAGCAAGCAAAATCATATATGAGAGC 633
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValaAsnGlyValaAspGlyTyrAspTyrSer 200
 Db 634 TATGGGAGTATTTGAGAGAGACATATGAAATGAGGATGATGCTATGATGATGATGATGATG 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluAspValGluAspValGluAspValGluAspValGlu 720

Db 694 CCGGCCAGTGTGATTAAGATGGAACATACCTTTGAGAGATTAACCATTAATATGAA 753
 QY 221 HSLLEuHISAlaTyValArgAlaLysLeuMetAsnAlaTyProSerTyTyrIleSerPro 240
 Db 754 CATCTTCATGCGCTATGTGAGGGGCAAAAGTGAATGAAGCCATATCCCTATATACAGTCA 813
 QY 241 ILleGlyCysLeuProAlaHISLeuLeuGlyAspMetTyrGlyArgPheTyrHisLeu 260
 Db 814 ATTTGATGCGCTCCCTCTCATTTGCTTGGTATATGTGGGATATTTGGCAAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 874 TACTCTTTCAGATTCCTCTTGGACAGAAACCAATAGATGATGATGCAATGGTG 933
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGlnAlaGluLysPheLeuValSerVal 300
 Db 934 GACCAGGCGCTGGATGCACAGAGAAATTCAGAGAGCCGAGAAAGTCTTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGlnAsnSerMetLeuThrAspProGlyAsn 320
 Db 994 GGTCTTCTTAATATGACTCAAGATTCCTGGAAATTTCCATGCTAAGCCAGCAAGAAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 Db 1054 GTTCAGAAAGCAGCTGCCATCCACAGCTTGGGACCTGGGAAAGGCGCACTTCAGAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyHis 360
 Db 1114 CTATATGTCACAAAGGTGACAAATGACAGACTTCTGACAGCTCATCATGATGAGGGGCAT 1173
 QY 361 ILleGlnTyAspMetAlaTyValAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1174 ATCCAGATGATGATGGCATATGCTGCACAACTTTTCTGCTAAGAAATGAGCTAATGAA 1233
 QY 381 GlyPheHisGlnAlaValGlnIleMetSerLeuSerIleAlaThrProLysHisLeu 400
 Db 1234 GGATTCATGAAAGCTGTGGGAATATGATGCTTCTGACAGCCATTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGlnIleThrGlnIleAsnPhe 420
 Db 1294 AAATCCATGTGCTCTGTCACCCGATTTTCAAGAAACAAATGAAACAGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheTyrMetLeuGlnLys 440
 Db 1354 CTGCTCAAAACAGACACACATGTTGGACCTGTCATTTACTTACATGATGAGAAAG 1413
 QY 441 TrpArgTyrMetValPheLysGlyGlnIleProLysAspGlnIleThrMetLysTyrTrp 460
 Db 1414 TGGAGGTGATGCTCTTAAAGGGGAAATTCACCAAGACCGATGATGAAAGTGTGG 1473
 QY 461 GluMetLysArgGlnIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 Db 1474 GAGATGAAGCAGAGATAGTGGGGTGTGGACCTGACCCCTGATGAAACCTATCTGT 1533
 QY 481 AspProIleSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 Db 1534 GACCCCGCATCTGCTTCATGTTCTGATGATATTCATTCATTCATTCATTAACAAAG 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnIleValLeuLysGlnAlaAlaLysHisGlnGlyPro 520
 Db 1594 ACCCTTACCAATTCATTCAGAAAGCAGCTTGTCAAGAGCAATTAACATGAAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPhe----- 537
 Db 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAAGTGTG- GTAAGAAAT 1712
 QY 537 ----- 537
 Db 1713 ACCCTAAATGTGAACCTCTCTAGTATTCAGTATTCATTCATTCATTCATTCATTCAT 1772
 QY 537 ----- 537
 Db 1773 TATTTGATTTCTTTGTTCTAAAGAAATTTTATGGCTCAAAATGTCTCATTTACAA 1832

QY 537 ----- 537
 Db 1833 ACCAAACATTTAATTTGTGTGTCAGACAGAAACCTAGAACATACAAATTTGGTGGGCA 1892
 QY 537 ----- 537
 Db 1893 CCTTTTCTCCCTATCATTAACTAAGCCCTCTCTCTGATTAATGGAAGAAAGAGCG 1952
 QY 537 ----- 537
 Db 1953 GTTAGGTGGAATATATCTGTATATGATTCCTTTTCTTATCTGACAGAAATTT 2012
 QY 537 ----- 537
 Db 2013 AGCCAAGTCAAGAGAGAAACCATAGATCATGATGAATATATGATCATCTGGAACC 2072
 QY 538 ----- 538
 Db 2073 CCTCAAAAGCCCTGAGACCCCTTTTGTGTAGCAATATATGAGGCTTGGAAATCA 2132
 QY 546 GluProTyrThrLeuAlaLeuGlnLysAsnValAlaGlyAlaLysAsnMetAsnValArgPro 565
 Db 2133 GAACCTGAGACCTTGCATGATGGAATATGTGTAGAGC- MAGAATGATGTAAGGCA 2191
 QY 566 LeuLeuAsnTyrPheGlnProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPhe 585
 Db 2192 CTGCTCAACTACTTGTAGCCCTTATTTACCTGGCTGAAAGACCAAGAACAAATTCCTTT 2251
 QY 586 ValGlyTyrSerThrAspTyrSerProTyrAlaAsp- GlnSerIleLysValArgIleSer 605
 Db 2252 GTGGAGTGGATGACCGAGCTGGAGTGCATATGACAGCCAAAGCATCAAGAGAGATGAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaArgIleThrPheLysValLysAsnGlnMetIle 625
 Db 2312 CCTAAATATGAGCTTGTGAGATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2371
 QY 625 eArgSerSerValAlaTyValAlaMetArgIleThrPheLysValLysAsnGlnMetIle 645
 Db 2372 CCGATATCTGTGCTATGCTATGAGCAGACGACTTTTAAAGTAAAAATCAGATGAT 2431
 QY 645 eLeuPheGlyGlnLysAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 Db 2432 TCTTTTGGGAGAGAGATGTCGAGTGCATTTGAACCAACCAAACTCTTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleLeuProArgThrGlnValGlnLysAl 685
 Db 2492 CTTCGCTACCTGACCTTAAATGATGTGTATATCATCTTACAGCTGAAATGAAAGGCC 2551
 QY 685 aLeaArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 Db 2552 CATCAGAGATCTCCGGAGCGGATATCAATGATGCTTCGCTGAAATGACAAAGCCCTGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2612 GTTCTGGGGATACAGCAACACTTGACCTCTTACCAAGCCCTC 2657
 RESULT 15
 E39034 2262 bp DNA linear PAT 18-JUN-2001
 LOCUS MPROT15 polypeptide and MPROT15 polynucleotide.
 DEFINITION E39034
 ACCESSION E39034
 VERSION E39034.1 GI:13017696
 KEYWORDS JP 1999318472-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2262)
 AUTHORS Christopher D.S. and Nicola B.
 TITLES MPROT15 polypeptide and MPROT15 polynucleotide
 JOURNAL Patent: JP 1999318472-A 2 24-NOV-1999;
 SMTKLIN BECHAM CORP PUBLIC LTD CO


```

Db 1060 GGCCTTCTCTATGATCAAGATCTGGGCAAACTCTGCTGATGCGCAGAGAT 1119
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgLeu 340
Db 1120 GGCGGAAATGTTGCTGCCACCCACAGCTGGGATCTGGGACACGGAGACTTCAGATC 1179
QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisGluMetGlyHis 360
Db 1180 AAGATGTGTCAAAAGGTCACAAATGACAACTTCTGACACCCATCAGCATGGGACAC 1239
QY 361 IleGluTrpAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1240 ATCCAAATATGATGATGATATGCGCAGCAACCTTCTGCTAAGAAAGGAGCAATGAA 1299
QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1300 GGCTTCATGACCTGTGGAGAAATCATGCTTCTGACACTACCCCAAGCATCTG 1359
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsn 420
Db 1360 AAATTCATGTTGCTTCTGCGCATCCGATTTTCAAGAAAGATGCGCAACAGATAACTTC 1419
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGlyHis 440
Db 1420 CTACTGAACAGCATGATGACATGTTGGACACTACCTTCTTACTTACATGTTAGAGAG 1479
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnThrMetLysLysTrp 460
Db 1480 TGGAGGTGATGTTGCTTGGGGTGAATTCACCAAGACAGATGATGAAAAAGTGTG 1539
QY 461 GluMetLysArgLysIleValGlyValAlaGluProValProHisAspGluThrTrpCys 480
Db 1540 GAATTAAGCGGGAGATCGTTGGTGGGAGACCTCTGCTCATGATGATAAACATCTCT 1599
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpArg 500
Db 1600 GACCTGCAATCTGTTCCATGTTCTAATGATTAATCTATCTCATTCGATTAATTCACAA 1659
QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
Db 1660 ACCATTACCAATTCAGATTCAGAAAGCTCTTGTGCAAGCAGCTAAGATATATGTTCT 1719
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db 1720 CTGCAACAATGTCATCTCAAAATTCCTGAGGCGGACAGAGTGTCTCAAGATGCTG 1779
QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsn 560
Db 1780 AGCTTGAATTCAGAGCCCTGAGCAACAAAGCTTGGAAATGCTGTAGAGACCAAGAT 1839
QY 561 MetAsnValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGln 580
Db 1840 ATGAGATGTAACCACTGCTCAATTAATTCACCAACGTTGTTGACTGCTGAAAGAGCAG 1899
QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspGlnSerIle 600
Db 1900 AACAGAAATCTTTTGTGGGTGCAACACTGATGAGCCCATTAATGCGCAACCAAGCAT 1959
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsn 620
Db 1960 AAGATGAGATTAAGCTTAATTAATCACTCTTGAAGCTAATGATGATGAGCAACAAAC 2019
QY 621 GluMetTrpLeuPheArgSerSerValAlaTrpAlaMetAlaGlnTrpPheLeuLysVal 640
Db 2020 GAAATGTTCTGCTGCAATCTGTTGCAATGCGCATGAGAAACTATTTTCAATATTC 2079
QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db 2080 AAAAACACAGACAGTTCCTTTCTAGAGAAAGATGACAGAGTGCAGGATTTGAACCAAGA 2139
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680

```

```

Db 2140 GTCTCTCTTACTCTTGTTCACCTACCCACCCAAATGATGTCATCTTCTGAGAT 2199
QY 681 GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db 2200 GAAGTTGAAGATGCCATCGATGATGTCGGGGCCCATCAATGATGTTGCTGGCTGAAT 2259
QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGluProPro 720
Db 2260 GATTAACAGCTGTGAGTTCTGGGGATTCACCCAAACATTTGAGCCACCTTACAGCTCT 2319

RESULT 17
BC026801
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source
CDS

```

2739 bp mRNA linear ROD 07-NDG-2002
 Mus musculus, RIKEN cDNA 2010305L05 gene, clone MGC:25940
 IMAGE:4236529, mRNA, complete cds.

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2739)
 Strausberg, R.
 Direct Submission
 Submitted (04-Apr-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amgdbcm.tmc.edu
 Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Hale, S. M.,
 Yoon, V. S., Kowls, C. R., Lawrence, S., Martin, R. G., Muzny, D. M.,
 Richards, S., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
 Series: IRAC Plate: 31 Row: e Column: 4
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.
 location/Qualifiers
 1..2739
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/A"
 /clone="MGC:25940 IMAGE:4236529"
 /tissue_type="Kidney, normal, 5 month old male mouse."
 /clone_id="NCI_CGAP_Kid14"
 /lab_host="DH10B"
 /note="vector: pCMV-Sport6"
 176..2593
 /codon_start=1
 /product="RIKEN cDNA 2010305L05 gene"
 /protein_id="AAH26801.1"
 /db_xref="GI:20071164"
 /db_xref="LOCUSID:70008"
 /translation="MSSSWLLISLVAVTTAOSITERNATFLINFRNOAEDISYSS
 LASNTNTNTEENAOQMSAAKWSAFYEQSKTQSFSLQEIPTPIKRLQALQO
 SGSSALSDADNKKOLNLTNTMSTIYSTGKVCNPKNOECILBPGIDEIMASTDYS
 RLWAMEGSAEVRKOLPLIEEVYVLKNEKARANNVNDYDWRGYEAGAGAYVN
 RNOLIEDVERTFAEIKPLYLHAYVRKIMLDYPSYISPTGLPAHLIGDMGRWT
 NLYLVEFAOKRPIVDYDAMNMGDAEKIIOEAKRFPYSGVLPHMTGQFVANSULT
 EPDGRKVVCHPTAMDGHDFTIKKCTKYTMNPLTAHHEKGIOTDYAKAPPLT

DB 2090 GAAATGCTTCCTTCGATCATCTGTCATATGCGATAGAAATATTTTCAATATC 2149

QY 641 LysAsnGlnMetIleuPheGlyGluAspValArgValAlaAsnLeuProArg 660

DB 2150 AAAAACACACAGATTCCTTTCTAGAGAGATGACAGTACGATTTGAACACAGA 2209

QY 661 IleSerPheAsnPhePheValIleAlaProLysAsnValSerAspIleIleProArgThr 680

DB 2210 GTCTCTCTCTACTCTTCTGTCACCTACCCCAAAATGTCGTGATGATCTCTGAGAGT 2269

QY 681 GluValGluLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsn 700

DB 2270 GAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2329

QY 701 AspAsnSerIleuGluPheLeuGlyIleGlnProThrIleuGlyProAsnGlnPro 720

DB 2330 GATAACAGCGCTGAGATTTCTGGGATTCACCCACACTTGAGCCACTTACAGGCTCT 2389

RESULT 18

AB053181

LOCUS 2760 bp mRNA linear ROD 03-APR-2001

DEFINITION Mus musculus ACE2 mRNA for angiotensin-converting enzyme-related

ACCESSION AB053181

VERSION AB053181.1 GI:13517092

KEYWORDS

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

1 (sites)

Komatsu, T., Sugano, S., Imai, J., Suzuki, Y., Hanaoka, K., Ymada, Y., Hida, M., Tanigami, A. and Muroi, S.

Molecular cloning, mRNA expression, and chromosomal localization of mouse Angiotensin-converting Enzyme-Related Carboxypeptidase

Unpublished

2 (bases 1 to 2760)

Komatsu, T., Sugano, S. and Suzuki, Y.

Direct Submission

Submitted (04-JAN-2001) Takami Komatsu, The Institute of medical science, University of Tokyo, Laboratory of Genome Structure Analysis, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan (E-mail: komatsu@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5283 (ex. 75283), Fax: 81-3-5449-5416)

FEATURES

source

1..2760

/organism="Mus musculus"

/db_xref="taxon:10090"

1..2760

/gene="ACE2"

187..2583

/gene="ACE2"

/codon_start=1

/product="angiotensin-converting enzyme-related

carboxypeptidase"

/protein_id="BAB40431.1"

/db_xref="GI:13517093"

/translation="MSSSSWLISLVAVTAASTLENAKTFLLNNFNOEADLSYQSS

LASVNYNTITEENAKQKSEAAKMSAEQSTAEQSTAEQSTAEQSTAEQSTAEQSTAEQ

SSSALSDKKNKNTILNTMSTYSGKVNPRNPCCILFEGIDEMIAKTDYNS

RLMMEGRARVQKOLRPLDEYVILKREMARANNYNGYRGADITABEADGENTN

RNQLIEDVERTFEAKPLIEHLHAYVRKRLMDYPSYISPTGCLPAHLIGMWGRFT

NLYPLTFEAPKPNIDVDAMNMGQWDERLFOAEKFEVSGLPHMFGWAMSMIT

EPADCKRVCHPTAMDGHDFRIKMCRTVMDNPLRAHHEMGHIOYDMAYARPELT

RNGANEGRHEAVGEIMLSAATPKHLSIGILPSDFQDSETEINFLKQALITVGLT

PEYIALEKRMWAVEGELSPKQMKKMKWEMKEKEIYGVPELPDETCDKALSTHVSN

DYSFIRYRTIYOFQOEALCOAKYVNGSLHKDISNSTEAGOKILKMSLSGSEPR

TKALENVYARNMDYKPLNFQPLPMLKQKNNSRYGNMTSPYADSDIKRISL

KSLGANAATENTNMELFNSVAYAMKRTYSLIKNIVPLLEDVAVRSDIKRVSF

FEVTSPOVNSVIRSEVEDAIRMSRGINDVFGINDSLFELIHPLLEPPYQPPVAT

IMLIIFGVVALVYVYIILITVGIKGRKKKRETKREESPDSTGHWKRRKROCRIPKO

BASE COUNT 830 a 585 c 628 g 717 t

ORIGIN

Alignment Scores:

Pred. No.: 3.98e-272 Length: 2760

Score: 3251.00 Matches: 598

Percent Similarity: 89.86% Conservative: 49

Best Local Similarity: 83.06% Mismatches: 73

Query Match: 84.03% Indels: 0

DB: 10 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AB053181 (1-2760)

QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20

DB 241 TCCCTCACCGAGAAATGCCAACATTTTAAACAACATTATACAGAACTGAAGAC 300

QY 21 LeuPheTyrglnSerSerLeuAlaSerTrpAsnTyraAsnThrAsnIleThrGluAsn 40

DB 301 CTGTCTTATCAAAAGTTCACTTGTCTTGGAAATTAATACTAATTAATTAATTAATTAAT 360

QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerLalaPheLeuysGluGlnSerThr 60

DB 361 GCCCAAAAGATGATGAGGCTGACGCCAAATGCTCTTTATTAAGAAAGCTCTAAG 420

QY 61 LeuAlaGlnMetTyrglnProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

DB 421 ACTGCCAAAGTTTCTCAGTACAGAAATTCAGACTCCGATTCATCAAGCTCAACTACAG 480

QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100

DB 481 GCCCTTACGAAAGTGGGTCTTCAGCTCTCAGCAAGCAAGAAACAGTGAACACA 540

QY 101 IleLeuAsnThrMetSerThrIleTyrglnSerThrGlyLysValLysAsnProAspAsnPro 120

DB 541 ATTGTGACACCATGACACCATTTACAGTGGAAATTTGCAACCCCAAGAACCCA 600

QY 121 GlnGlnCysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrgln 140

DB 601 CAAGAACCTTATTAATCTGAGCCAGATGAGAAATATGCGACAAACACAGACACTAC 660

QY 141 AsnGlnArgLeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyGlnLeuArgPro 160

DB 661 AACCTTACGAGCTCTGAGTGGAGGCTGAGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 720

QY 161 LeuTyrglnGluLysValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrglnAsp 180

DB 721 TTGTATGAAGAGTGTGTGCTCTGAAAACGAGATGCGAAGACAAATTTATTAACGAC 780

QY 181 TyrglnAspTyrglnTrpArgGlyAspTyrglnValAsnGlyValAspGlyTyrglnAspTyrgln 200

DB 781 TATGGGATTAATGAGAGGAGGACTATGAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

QY 201 ArgGlnGlnLeuIleGluAspValGlnHisThrPheGlnGlnIleLysProLeuTyrgln 220

DB 841 CGTACCAAGTTGATGAAGATGAGAAACGATCTTCCAGACATCAAGCATGTATGAG 900

QY 221 HisLeuHisAlaTyrglnValArgAlaLysLeuMetAsnAlaTyrglnProSerTyrglnIleSerPro 240

DB 901 CATTTTATGCTATGCTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960

QY 241 IleGlyLysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260

DB 961 ACTGAGATGCTCTCCCTGCGCATTTGCTGTGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020

QY 261 TyrSerLeuThrValLysProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280

DB 1021 TACCTTTGAGCTGTCTCTTGCACAGAAACCAACATTAATGTATGTATGAGAAAGGAGG 1080

QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerVal 300

DB 1081 AATCAGGCGCTGGATGAGAAAGATTTTCAAGAGGACAGAAATTTCTTGTCTGTT 1140

QY 301 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
 |||||.....
 Db 1141 GGCTCTCATATGACTCAAGATTCCTGGCAAACTGATGCTGACTGACAGAGAT 1200
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlySerGlyAspPheArgIle 340

 Db 1201 GGCGGAAAGTGTCTCCACCCACAGCTGGATCTGGACACGAGAGCTTCAAAATC 1260
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360

 Db 1261 AAGATGTGTACAAAGGTCACAAATGACAACTTCTTGACACCCATCAGAGATGGAGAC 1320
 QY 361 IleGlnTrpAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380

 Db 1321 ATCCATATGACATGATGATGACGACGACCACTTCTGCTAGAAAGGAGCAATGAA 1380
 QY 381 GlyPheHisGluAlaValIleGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400

 Db 1381 GGGTTCATGAAAGCTGTGGAAATCATGCTTCTGACCTACCCCAAGCATCTG 1440
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGlnThrGluIleAsnPhe 420

 Db 1441 AAATCATGTGCTCTTCCGATCGATTCCTCAAGAAATGACGAAACAGAGATTAACCTTC 1500
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyMetLeuGlyLys 440

 Db 1501 CTACTGAAACAGGATGACAAATGTGTGAAACACTACCGTTTACTTCAATGTTAGAGAG 1560
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnThrPheLysLysTrpTrp 460

 Db 1561 TGGAGGTGATGCTCTTGGGGTGGAAATTCCTCAAGAGAGATGAGAGAAAGTGGTGG 1620
 QY 461 GluMetLysArgGlnIleValIleValIleProValProHisAspGlnThrTyCys 480

 Db 1621 GAGATGAAAGGGGAGATCGTGTGTGGAGGCTCTGCTCATGATGAACATCACTCT 1680
 QY 481 AspProLysSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500

 Db 1681 GACCTCATCTCTGCTCCATGTTCTTAAGATATCAATTCATTCATATTCACACAGG 1740
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuCysGlnIleAlaLysHisGluGlyPro 520

 Db 1741 ACCATTACCAATTCAGATTTCAGAAAGCTCTTGTACAGAGCTTAAGTAAATGCTTCT 1800
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540

 Db 1801 CTCGACAAATGTGACATCTCAATTCACAGAGCTGGGACGAAAGTGTCTCAAGATCTG 1860
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsn 560

 Db 1861 AGTCTTGGAAATTCAGAGCCCTGAGCCAAAGCCCTTGGAAATGTGGTAGAGAGAGAT 1920
 QY 561 MetAsnValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGln 580

 Db 1921 ATGATGTGAAACACACGCTCAATTAATTCACACCGTGTGTGACTGTGCTGAAAGACAG 1980
 QY 581 AsnLysAsnSerPheValIleTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600

 Db 1981 AACAGAAATCTCTTGTGGGGTGGACACAGTGAAGGAGCCCATATGCCGACCAAGCAT 2040
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyLysAspLysAlaTrpGluTrpAsnAspAsn 620

 Db 2041 AAATGTGAGGATTAACCTTAATTCAGCTTGGAGCTTAATTCATATGATGAGCAACAC 2100
 QY 621 GluMetTrpLeuPheAspSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysVal 640

 Db 2101 GAATGTCTCCGTTCGATCATCTGTGCAATGCCATGAGAAAGTATTTTTCATATATC 2160
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArg 660

 Db 2161 AAAAACCAGACAGTCTCTTCTAGAGAGAGATGATGAGATGATTTGAAACCAAGA 2220
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680

Db 2221 GTCTCTTCTACTTCTTGTCTACCTACCCCAAAATGTGTGATGATCTCTAGAGT 2280
 QY 681 GluValGlnLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsn 700

 Db 2281 GAAGTTGAAGATGCCATGAGATGTCTGGGGCCCATCAATGATGTCTTGGCCGTGAT 2340
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720

 Db 2341 GATTAACAGCTGTGAGTTCTGTGGGATTAACCAACACTTGAGGACCATTTACAGAGCTCT 2400
 RESULT 19
 AX047765
 LOCUS AX047765 2638 bp DNA linear PAT 15-DEC-2000
 DEFINITION Sequence 8 from Patent WO0070032.
 ACCESSION AX047765
 VERSION AX047765.1 GI:11876771
 KEYWORDS
 SOURCE
 ORGANISM
 Mus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
 Zacez: a human metalloenzyme
 Patent: WO 0070032-A 8 23-NOV-2000;
 Zymogenetics, Inc. (US)
 Location/Qualifiers
 1..2638
 /organism="Mus sp."
 /db_xref="taxon:10095"
 106..2523
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC18993.1"
 /db_xref="GI:11876772"
 /translation="MSSSWLLSLVAVTTRQSLTEENAKFLNNFNOEADLSYSS
 LASWNTNTEENAKMSEAAKSAVEEQSPKQAFSLQETQPIKRLQALQO
 SSSLSLSDKNKQNTIINTMSTIYSTKVNPRPOCDLLEGLDEIMATSTDYNS
 RLWAMEGMAEYKQRLPLEYERYVLLKEMRANRYNDYGDYKADYAEADQYNN
 RQULEDERYFAEIKPLYEHLHAYRKRLMDYYSYISPGCIPAILHGLMGMRFT
 NLYPLTPFAQPKPNDIYDAMNQCWDRLRFOAEKRFYVGIPLHMGCRWASMT
 EPADGRKYVCHPADLGHGDFRIKCTKYTMDFLRHHEMGIQDMAYARQPELL
 RKGANEGEHVAEGLMSLAATPKHLKIGLPSDFQDSLETINFLKQALTVGL
 PPTVLMLEKRWAVFGEIPEKDMKMKWEMREIVGVEPDRBETCYDPSLSEVSN
 DYSFTRYRTYFOFQALCOAKYNGSLHKDISNSTAGOKLKLMSLGNSEPN
 TALENVGARMDYKPLNTYFOPLDMYKQNRNRSFGMTNEMSPYDOSIKRISL
 KSLGANAVENTNMNEMPLFRSSVAAMARKSSIIKNQVPLLEDYRVSQDKPRVSY
 FVYTSFQNSDYTPRSEVDAIRMSRGRLNDYFGINDSLFLGTHPTLEPYQPT
 IWLIFGVVMAIIVVGIILITVGIKGRKKRNETKREBNPYDSMDIGKESNAGFONS
 DDAQTSF"
 BASE COUNT 797 a 557 c 615 g 669 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,37e-270 Length: 2638
 Score: 3233.00 Matches: 596
 Percent Similarity: 89.58% Conservative: 49
 Best Local Similarity: 82.78% Mismatches: 75
 Query Match: 83.56% Indels: 0
 DB: 6 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AX047765 (1-2638)
 QY 1 SerThrIleGlnGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 |||.....
 Db 160 TCCCTCACCGAGGAATATCCCAAGCATTTTAAACACTTAAACAGAGAGCTTAACAC 219
 QY 21 LeuPheTrpGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGluGluAsn 40
 |||.....
 Db 220 CTGTCTTATCAAAAGTCACTTCTGCTTGTGAATTAATTAATCAATTAATCAAGAAAT 279
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60

Db	280	GCCCAAAAGATAGTAGGCTGCAGCCCAATGGTCTGCTTATATGAAAGACGTATAG	339
OY	61	LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValIleLysLeuGlnLeuGln	80
Db	340	ACTGCCCAAGTTCTTCACTACACAGAAATCCAGACTCGATCATCAAGGCTCAACTACAG	399
OY	81	AlaLeuGlnGlnIleAsnGlySerSerValIleSerGluIleAspLysSerLysArgLeuAsnThr	100
Db	400	GCCCTTCACAGAAAGTGGCTTCAGACACTCTCAGACGACAAAGAACAAACAGTTGAACACA	459
OY	101	IleIleAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro	120
Db	460	ATTCTGAACCATATGACACCACTTATACATCTACTGTGAAAAGTTTGCAACCCAGGAACCCA	519
OY	121	GlnGlnCysLeuLeuLeuGlnIleProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr	140
Db	520	CAGAATGCTTATTACTTGAGCCAGAGATTGGATGAAATATATGCGCAAGACACAGACTAC	579
OY	141	AsnGlnArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro	160
Db	560	AACTCTAGGCTCTGGGCATGGAGGGCTGGAGCGCTGAGGTTGGCAAGCAGCTGGAGCGCC	639
OY	161	LeuTyrGlnGlnIleTyrValValIleLysAsnGlnIleMetAlaAsnHisTyrGluAsp	180
Db	640	TTGTATGAAGAGTATGTGCTCTGAAAAACGAGATGGCAGAGGCCAAACATTTATACAGC	699
OY	181	TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer	200
Db	700	TATGGGGATTTATTGGAGAGGGGACACTATGAACGACAGAGGACGAGATGGGCTACACATATAC	759
OY	201	ArgGlyGlnLeuIleGluAspValGlnHisThrProGlnGlnIleLysProLeuTyrGlu	220
Db	760	CGTAAACGATTGATTGAAGATGTAAACGTCATCTTCGACAGATCAAGCCATTGTATGAG	819
OY	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrTrpLeuPro	240
Db	820	CATCTTCATGCTATGTAGTAGAGGAGGAAGTTGATGATACCTACCTTCCTCAACTACGCCCC	879
OY	241	IleGlyCysLeuProAlaHisLeuLeuLysAspMetTrpGlyArgProThrTrpAsnLeu	260
Db	880	ACTGGATCCCTCCCGCCCAATTTGGTGGTATATGAGGGAGATTTTGGCAAAATGTG	939
OY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	940	TACCCCTTGACTGTCCCTTGGCACAGAAACCAACATAGATGTTACTGATGCAATGACG	999
OY	281	AspGlnAlaTrpAspAlaGlnArgTlePheLysGluAlaGluLysPhePheValSerVal	300
Db	1000	AATCAGGGCTGGATGTCAGAAAGGATATTTCACAGGACGAGAAATTTCTTGTTCTGTT	1059
OY	301	GlyLeuProAsnMetThrGlnGlyPheThrProLysAsnSerMetLeuThrAspProGlyAsn	320
Db	1060	GGCCTCTCTCATATGACTCAAGAGATTTGGGCAAACTGTATAGCTACATGAGCCACCAAT	1119
OY	321	ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTle	340
Db	1120	GGCCGGAAGATGTCTGCCACCCACACACTTGGGATCTGGGACACGAGACTTCGAAATC	1179
OY	341	LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis	360
Db	1180	AAGATGTGTAAAGGTACATATGACAACTTCTTGACAGGCCATCAAGAGATGGGACAC	1239
OY	361	IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGlu	380
Db	1240	ATCCAAATATGACATGGCATATGCCACGAGCAACCTTCCCTGTAAGAAACGAGCCATATA	1299
OY	381	GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu	400
Db	1300	GGGTTTCATGAAGCTGTGGAAATCATGTCACTTCTCGACGCTACCCCAACGACATGTG	1359
OY	401	LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIlePhe	420

LOCUS	DEFINITION	VERSION	KEYWORDS	ORGANISM	REFERENCE
AX047760	Sequence 3 from Patent WO0070032.	AX047760	AX047760.1 GI:11876767	synthetic construct. artificial sequences. 1 (bases 1 to 2415)	
1366	AAATGCATTGGCTCTTGGCAATCCGATTTTCAAGAAGATAGCGAAACAGAGATAAATCTC	1411			
421	LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLys	440			
1420	CTACTGAAACAGCGCAATTCACAAATTTGTGGACACATACCGTTACTTACTGTTACAGAAG	1475			
441	TPAAGTTPMeValPheLysGlyGluIleProLysAspGlnTyrMetLysLysTyrP	460			
1480	TGGAGGTGATGTCTTTCGGGGTGAATAATCCCAAGAGCACTGGATGAAAAAGGGTGG	1538			
461	GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys	480			
1540	GAGATGAAGCGGAGAGATCGTTGTGGTGGAGCGCTCGCTCGATGAACAAATACGT	1599			
481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg	500			
1600	GACCTCGATCTCTTCCATGTTTCTAATGATTTCTATTCATTCATTCGATATTAACAAG	1655			
501	ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaIleLysHisGlnGlyPro	520			
1660	ACCATTTCACATTCACATTCACAAAGCTCTTTTCAAGCAGCTTAAGATTAATGGTCT	1719			
521	LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnLeu	540			
1720	CTGCAACAAATGTGACATCTCAAAATCCACTGAAGCTGGGAGAAATGCTCTCAAGATGCTG	1779			
541	ArgLeuGlyLysSerGluProThrPheLeuAlaLeuGlnAsnValValGlyValAlaLysAsn	560			
1780	AGCTCTGGAATTCAGACGCTCGACCCAGACGCTTGGAATAATGTGTGAGCAAGGAT	1839			
561	MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTyrPheLysAspGln	580			
1840	ATGTGATGAACACCACTGCTCAATTAATCTCAACCGCTTGTTGACTGCTGAAAGAGCAG	1899			
581	AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle	600			
1900	AACGAAATTTCTTTGTGGGTGGAGAACCTGAATGAGACCCATATGCGACCAAAAGCATT	1959			
601	LysValAlaArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlyLTPAsnAspAsn	620			
1960	AAAGTGAGGATTAAGCCTTAATAATCAGCTTGTGAGCTAATGCATATGATGACCAACACAC	2019			
621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640			
2020	GAAATGGTCCGTGTCGATCATCTGTTGCCAATGCGATGACGAAAGATTTCTTCAATATC	2079			
641	LysAsnGluMetIleLeuPheGlyGluGlnLysPValArgValAlaAsnLeuLysProArg	660			
2080	AAAAACCGACGAGATTCCTTTCTAGAGGAAGATGTACGAGTGAAGTATTAAACCAAGA	2139			
661	IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr	680			
2140	GTCTCCCTCTCTCTCTCTGTCACCTCAACCCCAAAATGTGTGATGTCAATCTCAAGAGT	2199			
681	GlyValGlyLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700			
2200	GAACTTGAAAGTGCATCAGCATATGTCTCGGGGCGCGCATCAATGATGTCTTGGCCGTAAT	2259			
701	AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro	720			
2260	GATAACAGACCTCGAGATTCTTGTGGGATTCACCAACACTTGAGCCACCTTACCAACCTCCT	2319			

RESULT	20
AX047760	
LOCUS	AX047760
DEFINITION	Sequence 3 from Patent WO0070032.
ACCESSION	AX047760
VERSION	AX047760.1 GI:11876767
KEYWORDS	.
SOURCE	synthetic construct.
ORGANISM	synthetic construct
REFERENCE	artificial sequences. 1 (bases 1 to 2415)
DNA	linear
PAT	15-DEC-2000

AUTHORS Piddington, C.S., Petrie, C.R., Shoemaker, K.E. and Bishop, P.D.
 TITLE Zacc2: a human metalloenzyme
 JOURNAL Patent: WO 00/0032-A 3 23-NOV-2000;
 ZymoGenetics, Inc. (US)

FEATURES
 source
 1. .2415
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="This degenerate sequence encodes the amino acid
 sequence of SEQ ID NO:2."
 1. .2415

misc_feature
 /note="n = A,T,C or G"

BASE COUNT 494 a 218 c 398 g 335 t 970 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.3e-265 Length: 2415
 Score: 3175.00 Matches: 578
 Percent Similarity: 80.39% Conservative: 0
 Best Local Similarity: 80.39% Mismatches: 141
 Query Match: 82.06% Indels: 0
 DB: 6 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AX047760 (1-2415)

Qy 2 ThrileguglunalaalysrThrPheLeuAspLysPheAsnHsGluAlaGluAspLeu 21
 Db 58 ACNATHGARGARCARCNCNARACNTTYTNGATAARTTAAATCAATGARGCNGARGAYTN 117
 Qy 22 PheTyrGlnSerSerLeuAlaSerTyrPAsnTyrAsnThrAsnIleThrGluAsnVal 41
 Db 118 TTYTAYCARWSNMSNTYTCNMSWTGGAATTAAYAAACNAAYATHCNGARGARAAYGTN 177
 Qy 42 GluAsnMetAsnAlaGluAspLysTyrSerAlaPheLeuLysGluGlnSerThrLeu 61
 Db 178 CARATATGATATATGCGNGNGAYAAATGCGMSNCNTTYTNAARAKARCARMSNACNTN 237
 Qy 62 AlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAla 81
 Db 238 GCNCARATGTAAACNTYNCARAGATTCARAAYTNAACNTNARNTNARNTNARNTNAR 297
 Qy 82 LeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIle 101
 Db 298 YTNRCARARAYGGMSNMSNGNTYTNMSNGARATTAARMSNARMSNARNTNAAATACNATH 357
 Qy 102 LeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnProGln 121
 Db 358 YTNAAAYACNATGMSNACNATHTAATWMSNACNGNARNTNGYAAYCCNGAYCCNGCAR 417
 Qy 122 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 141
 Db 418 GARGGYTNTYNTYNTGNCRCNGCYTNAAGARATHTATGCGNAAWMSNTYNGAYTAAAY 477
 Qy 142 GluArgLeuTyrPalaTyrPgluSerTyrPArgSerGluValGlyLysGlnLeuAspProLeu 161
 Db 478 GARGNCTNTTGGCCTGCGGARGMSWTGCGMSNKSARKTNGSNAARCARARNTNMGCCYTN 537
 Qy 162 TyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHsIleTyrGluAspTyr 181
 Db 538 TAYGARGARATYGTNGTNTYTAARAAYAGARATGCGNMGNCNAAAYCATYTAGARGATYAY 597
 Qy 182 GlyAspTyrTyrPArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArg 201
 Db 598 GGNAGYTAATYGTGNGGNGATYTAGARGTNAAYGCGNTGAYGNTAYGATTAATWMSNGN 657
 Qy 202 GlyLysLeuIleGluAspValGluHsIleThrPheGluGluIleLysProLeuTyrGluHsI 221
 Db 658 GGNCAATYNTATHGARGATGTGARGATCAACNTTTCARARATTHARCCNTYNTATGARGAY 717
 Qy 222 LeuHsIleTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIle 241
 Db 718 YTNCAAYGCTAYGTGMSGNCNAAATYTNATGAAYCCNTAYCCNMSNTAYATHTMSNCCNATH 777

Qy 242 GlyCysLeuProAlaHsIleLeuGlyAspMetTyrPgluArgPheTyrThrAsnLeuTyr 261
 Db 778 GGNCTYTNCCNCGNCATYTTNTNGNCAATATGCGGCMGNTTYTGACNAAAYTTNTAY 837
 Qy 262 SerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAsp 281
 Db 838 WSNYTNACNCGNCCWTYTGNCARARCNAAATHTGAYGTNMGNGAYCGCATGTGTGAY 897
 Qy 282 GlnAlaTyrPalaGluArgIlePheLysGluAlaGluLysPhePheValSerValGly 301
 Db 898 CARGCNTGGAGCGCACMGATHTTAAARARCGNARARNTTYTGTWMSNGTNGN 957
 Qy 302 LeuProAsnMetThrGlnGlyPheTyrPgluAsnSerMetLeuThrAspProGlyAsnVal 321
 Db 958 YTNCCNAAATATGACNCCARGATTTTGGARAAWMSNATGTNACNGAYCCGNCNAAAYGTN 1017
 Qy 322 GlnLysAlaValCysHsIleProThrAlaTyrPAspLeuGlyLysGlyAspPheArgIleLeu 341
 Db 1018 CARAARGCNGTNTGCAVCNCCNCGNTGGAYTNGGNAARGGATTYTGMNATHYTN 1077
 Qy 342 MetCysThrLysValThrMetAspPheLeuThrAlaHsIleGluMetGlyHsIle 361
 Db 1078 ATGTGTACNAAARGTNAACATGATGATYTTTNCNCGNCAYCAVAGARATGCGNCAVATH 1137
 Qy 362 GlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGly 381
 Db 1138 CARTAYGATATGCTTAAVCGNCCNARCCNTTYTNTYNTNMGNAYGGGNCNAAAYGARGCN 1197
 Qy 382 PheHsIleGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHsIleLys 401
 Db 1198 TTYCAAGAGCNGTNGGARGATHTGMSNTTMSNCGNCCNACNCAACNATYTNAR 1257
 Qy 402 SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 421
 Db 1258 WGNATHGNGYTYTNTWMSNCCNCAATYTCARAGARGAAYAGARACNGARATTAATYTYTN 1317
 Qy 422 LeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTyr 441
 Db 1318 YTNAAACARCGNTYTNACNHTGTCNACNTYNTCCNTTYACATATATGTNGARARATGG 1377
 Qy 442 ArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysTyrTyrPglu 461
 Db 1378 MNTGATGTGTTTAAARGNARGARATHTCCNARGAAYCARTGATCAARAARATGGTGGAR 1437
 Qy 462 MetLysArgGluIleValGlyValGluProValProHsAspGluThrTyrCysAsp 481
 Db 1438 ATCAARMGARGATHTGTCNCGTNGTCNCCNCAAGAGARACNTAYTGAY 1497
 Qy 482 ProAlaSerLeuPheHsIleValSerAspTyrSerPheIleArgTyrThrArgThr 501
 Db 1498 CCNCGMSNTYTTTCATGTMSNAAATGATYATWMSNTYATHTMGATATTAACNMGNCN 1557
 Qy 502 LeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHsIleGluIleProLeu 521
 Db 1558 YTNATYCATRTTCARTTCARAGRCNTYNTYTCARGCNGCNAACAGARGCNCNTYN 1617
 Qy 522 HisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArg 541
 Db 1618 CAATAATGTGATATHTSNAAYMSNACNGARGCNGGCAAAATYTTTAATATGTGTMGN 1677
 Qy 542 LeuGlyLysSerGluProTyrThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMet 561
 Db 1678 YTNMGNAARMSNARGCNTGACNTYNTNGNTNGARAATGINTGNGNCAARAAYTNG 1737
 Qy 562 AsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrLeuLysAspGlnAsn 581
 Db 1738 AAYGTMGCCNTYNTNAAATYATYTAGRCNNTYTTACNTGTGTNAARAAYCARAY 1797
 Qy 582 LysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerLysLys 601
 Db 1798 AARAATWMSNTTGTGNTGNGMSNACNGAYTGTGMSNCCATATGCGATCARMSNATTHAR 1857
 Qy 602 ValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPAsnAspAsnGlu 621

```
Db      1858 GTNMGNAHTHMSNTTNAARMSNGCNYTNGNGATTAARCCNTATGARTGGAAYGAAYGAR 1917
QY      622 MetTyrLeuphearSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLys 641
Db      1918 ATGTAAYTNTTYMGMSMSNGTNGCNTATGCNATGMCNCAATATTTTNTNNAARGTNAAR 1977
QY      642 AsnGlnMetIleLeupheGlyGluGluAspValArgValAlaAsnLeuLysProArgIle 661
Db      1978 AATCAATGATHTTNTTYGGNGARGAGATNGMGNTGNCNAAYTNAARCCNMGNAHTH 2037
QY      662 SerPheasnPheValThrAlaProLysAsnValSerAspIleIleProArgThrGlu 681
Db      2038 MSNTTYAATTTTTCGNAACNGCCNAAAPATGTNMSGATATHTAHCNMGNAACNGAR 2097
QY      682 ValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAsp 701
Db      2098 GTNGARARARCCNATHMGNAATGMSMGMSMGNAATHAAGAYGCNTTYMGNTTNAAYGAY 2157
QY      702 AsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2158 AATWSNTTNGARTTTTNGSNATHCARCCNACNTNGNCCNCAAYCARCCNCCN 2214
```

Search completed: May 26, 2003, 19:37:57
Job time : 5047 secs

100

100

100

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	3869	100.0	2418	21	AAZ59465	Human MPROT15 coding sequence #1.
2	3869	100.0	3334	22	AAZ59465	Human MPROT15 coding sequence #1.
3	3869	100.0	3336	21	AAZ59465	Human MPROT15 coding sequence #1.
4	3869	100.0	3336	22	AAZ59465	Human MPROT15 coding sequence #1.
5	3869	100.0	3336	24	AAZ59465	Human MPROT15 coding sequence #1.
6	3869	100.0	3336	24	AAZ59465	Human MPROT15 coding sequence #1.
7	3833	99.1	2920	22	AAZ59465	Human MPROT15 coding sequence #1.
8	3785	97.8	2911	22	AAZ59465	Human MPROT15 coding sequence #1.
9	3720	96.1	3732	22	AAZ59465	Human MPROT15 coding sequence #1.
10	3366.5	87.0	2638	21	AAZ59465	Human MPROT15 coding sequence #1.
11	3251	84.0	2638	22	AAZ59465	Human MPROT15 coding sequence #1.
12	3233	83.6	2638	22	AAZ59465	Human MPROT15 coding sequence #1.
13	3214	83.1	2350	24	AAZ59465	Human MPROT15 coding sequence #1.
14	3175	82.1	2415	22	AAZ59465	Human MPROT15 coding sequence #1.
15	2832	73.2	3474	22	AAZ59465	Human MPROT15 coding sequence #1.
16	2640	68.2	2415	22	AAZ59465	Human MPROT15 coding sequence #1.
17	1337	34.6	2477	12	AAZ59465	Human MPROT15 coding sequence #1.
18	1337	34.6	4020	21	AAZ59465	Human MPROT15 coding sequence #1.
19	1337	34.6	4022	24	AAZ59465	Human MPROT15 coding sequence #1.
20	1337	34.6	4024	11	AAZ59465	Human MPROT15 coding sequence #1.
21	1337	34.6	4024	20	AAZ59465	Human MPROT15 coding sequence #1.
22	1336	34.5	4020	19	AAZ59465	Human MPROT15 coding sequence #1.
23	1334	34.5	3939	22	AAZ59465	Human MPROT15 coding sequence #1.
24	1334	34.5	4563	22	AAZ59465	Human MPROT15 coding sequence #1.
25	1310	33.9	3942	20	AAZ59465	Human MPROT15 coding sequence #1.
26	1275	33.0	5005	22	AAZ59465	Human MPROT15 coding sequence #1.
27	1233.5	31.9	2568	24	AAZ59465	Human MPROT15 coding sequence #1.
28	1059	27.6	2089	23	AAZ59465	Human MPROT15 coding sequence #1.
29	1054	27.2	2074	16	AAZ59465	Human MPROT15 coding sequence #1.
30	1025.5	26.5	2450	23	AAZ59465	Human MPROT15 coding sequence #1.
31	959	24.8	9006	22	AAZ59465	Human MPROT15 coding sequence #1.
32	924.5	23.9	5632	23	AAZ59465	Human MPROT15 coding sequence #1.
33	919.5	23.8	5060	23	AAZ59465	Human MPROT15 coding sequence #1.
34	841	21.7	2379	23	AAZ59465	Human MPROT15 coding sequence #1.
35	721	18.6	2082	21	AAZ59465	Human MPROT15 coding sequence #1.
36	721	18.6	2082	22	AAZ59465	Human MPROT15 coding sequence #1.
37	721	18.6	2082	24	AAZ59465	Human MPROT15 coding sequence #1.
38	715.5	18.5	1395	22	AAZ59465	Human MPROT15 coding sequence #1.
39	687	17.8	446	24	AAZ59465	Human MPROT15 coding sequence #1.
40	597.5	15.4	42573	23	AAZ59465	Human MPROT15 coding sequence #1.
41	491	12.7	2025	23	AAZ59465	Human MPROT15 coding sequence #1.
42	458	11.8	1836	23	AAZ59465	Human MPROT15 coding sequence #1.
43	441	11.4	5116	23	AAZ59465	Human MPROT15 coding sequence #1.
44	431	11.1	4001	23	AAZ59465	Human MPROT15 coding sequence #1.
45	429	11.1	561	24	AAZ59465	Human MPROT15 coding sequence #1.

PD 24-NOV-1999.
 XX 22-JAN-1999; 99JP-0014949.
 XX 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX (SMIK) SMITHKLINE BEECHAM PLC.
 PA WPI: 2000-109268/10.
 DR P-PSDB; AAY67310.
 XX MPRO15 polypeptide and MPRO15 polynucleotides - useful for the
 PT treatment of hypertension, myocardial diseases, apoplexy, heart
 PT diseases, nervous denaturation, Alzheimer's disease etc.
 XX Claim 7; Page 14; 22pp; Japanese.
 PS
 XX This is the coding sequence of human MPRO15. The MPRO15 polynucleotide
 CC and polypeptide sequences can be used for the treatment of hypertension,
 CC myocardial diseases, apoplexy, heart diseases, nervous denaturation,
 CC Alzheimer's disease and diseases related to the processing of peptide
 CC hormones and cytokines.
 XX
 SQ Sequence 2418 BP; 744 A; 484 C; 555 G; 635 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 2418
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AAY59465 (1-2418)

QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 55 TCACACATTTGAGGAAAGGACGACATTTTGGACAAAGTTTACACGAAAGCCGAGAC 114
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnTyrAsnHisGluGlnAla 40
 DB 115 CTGTCATATCAAGCTTCACCTTCTTGGATTTAATTAACCAATATTTACTGAGAGAAAT 174
 QY 41 ValGlnAsnMetAsnAlaGluAspLysPheSerAlaPheLeuLysGluGlnSerThr 60
 DB 175 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAAACAGTCCACA 234
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeu 80
 DB 235 CTTCGCCAAATGATTCACCTCAAGAAATTCAGAAATTCACAGCTTCAGCTGACG 294
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 295 GCTCTTCAGCAAAATGGGCTTCAGTCTCTCAGAAAGCAAGCAACGAGTTGAAACACA 354
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAspPro 120
 DB 355 ATTTCAATATCAAGAGACACATCTACAGTCTGGAAGTTTGTAAACCCAGATAATCCA 414
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
 DB 415 CAAGAATGCTTATTAATCTGACACAGGTTTGAATGAATTAATGCAACAGTTTGAATAC 474
 QY 141 AsnGluArgLeuThrPheAlaTyrPheLysSerThrArgSerGluValAlaGlyLysGlnLeuArgPro 160
 DB 475 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGGCAAGCAAGCTGAGGCCA 534
 QY 161 LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 535 TTATATGAAGATATGCTGCTTGAATAATAGATGGCAAGCAAAATCATTTAGAGAC 594
 QY 181 TyrGluAspTyrTyrPheGlyAspTyrGluValAlaAsnGluValAspGlyTyrAspTyrSer 200

DB 595 TATGGGATTTATGGAGAGACATATGAATTAATGGGATGATGGCTATGACTACAGC 654
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 655 CGCGGCCAGTGTATTCAGATGATGCAACATACCTTTGAAGATTAACCATTAATATGA 714
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 715 CATCTTCATGCTATGTAGGAGGCAAGTTGATGAATGCCCTATCTCTTATATCACTCCA 774
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 DB 775 ATTGATGAGCTCCGCTCATTTGCTGGTGTATGTGGGGAGATTTTGGCAAAATCTG 834
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 835 TACTCTTGCACAGTCCCTTGGACAGAAACCAACATAGATGTTACTGATGCAATGGTG 894
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 895 GACCAAGCCTGGGATCCACAGAGATATTCAGAGAGCCGAGAAATTTCTTGTATCTGT 954
 QY 301 GlyLeuProAsnMetThrGlnIlePheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 955 GGTCTTCATATATGACTCAAGAGATCTGGAAATTCATGCTTAACGAGCCAGAGAAAT 1014
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 DB 1015 GTTCACAAAGAGAGCTGCGCATCCACAGCTTGGAGCTGGGAGAGGAGCTTCAGAGATC 1074
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 DB 1075 CTTATGTGCACAAAGGTGACATGAGACGACTTCCTGACAGCTCATCATGATGGGGAT 1134
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1135 ATCCAGATGATATGATGATATGCTGACACAACTTTCTGCTAAGAAATGAGAGCTATGAA 1194
 QY 381 GlyPheHisGluAlaValAlaGlyIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1195 GATTCATGAGAGCTGTGGGAAACATGATGACATTTCTGCGACCCACACTTAAGCATTTA 1254
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnHe 420
 DB 1255 AATTCATGATGCTTCTGACCCGATTTCAAGAAAGCAATAGAAACGAAATTAACCTTC 1314
 QY 421 LeuLeuLysGlnAlaLeuThrIleValAlaGlyThrLeuProPheThrTyrMetLeuGlu 440
 DB 1315 CTGCTCAAAACAGACATCGATGTGTGGAGCTGCGCAATTTACTTACATGTATAGAGAAG 1374
 QY 441 TyrArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysTyrTyrTrp 460
 DB 1375 TGGAGGTGATGCTCTTTTAAAGGGAATTTCCCAAGACGATGATGAATAAGAGTGG 1434
 QY 461 GluMetLysArgGluIleValAlaGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1435 GAGATGAGAGAGATAGTGGGGGTGGGAACCTGTCGCCATGATGAAGAAACATAGCT 1494
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1495 GACCCGCGAGTCCTGTCATGTTCTTAATGATTCCTCATTCATTCGATATTAACAAGG 1554
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluIlePro 520
 DB 1555 ACCCTTACCAATTCAGTTTCAAGAGCACTTTGTCAAGCAAGCTTAACATGAAGCCCT 1614
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1615 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAATCTGTCAATGTGG 1674
 QY 541 ArgLeuGlyLysSerGluProThrPheLeuAlaLeuGluAsnValValGlyAlaLysAsn 560


```

Db 1675 AGGCTTGGAATTCAGAACCTTGACCTAGATTGGAAATGTTGTAGACCAAGAAC 1734
QY 561 MetasnlValArgProLeuLeuasnTyrrPheGluProLeuPheThrTyrrLeuLysAspGln 580
Db 1735 ATGAATGTAAAGCCACCTGCTCAACTTGTGAGCCCTTATTTACCTGGCTGGAAGCCAG 1794
QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTyrPseProTyrAlaAspGlnSerIle 600
Db 1795 AACCAAAATCTTTGTTGGATGGAGTACCGACTGGAGCCATATGACACCAAGCAATC 1854
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
Db 1855 AAAGTGAAGATTAAGCTTAATACAGCTCTTGAGAGATTAAGCATATGATGAGACGACAT 1914
QY 621 GluMetTyrrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrrPheLeuLysVal 640
Db 1915 GAAATGTACCTGCTTCGATCATCTGTGATATGATATGATGAGCAGACTTTTAAAGTA 1974
QY 641 LysAsnGluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db 1975 AAAAATCAGATATCTTTTGGGAGGAGATGCGCGAGTGGCTAATTTGAAACCAACA 2034
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
Db 2035 ATCTCCTTAATTTCTTTGTCTACTGACCTAAATATGCTCATATCTTCTAGACT 2094
QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db 2095 GAAGTTGAAAGAGCCATCAGATGATGTCGCCGAGCCGATCAATGATGCTTCCGTCGAT 2154
QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db 2155 GACAAACAGCTGAGATTTCTGGGGAGTACAGCCAAACATCTGACCTCAACAGCCCTT 2214
RESULT 2
AAC84366
ID AAC84366 standard; cDNA; 3334 BP.
XX
AC AAC84366;
XX
DT 19-MAR-2001 (first entry)
XX
DE Human Zace2 protein encoding cDNA.
XX
KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
KW ventricular systolic dysfunction; renal impairment; heart failure;
KW scleroderma renal crisis; atherosclerosis; antinflammatory; human;
KW antithrombotic; bradykinin inactivator; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 35..2452
FT /tag="a
FT /product="Zace2"
XX
XX MO200070032-A1.
XX
XX 23-NOV-2000.
XX
XX 03-MAY-2000; 2000MO-US11932.
XX
XX 13-MAY-1999; 9905-0311482.
XX
XX 27-AUG-1999; 9905-0384706.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
XX
XX WPI; 2001-025018/03.
XX
XX P-PSDB; AAB48095.
XX

```

```

PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
PT associated with inflammation such as arthritis and enterocolitis -
PS
XX Example 1; page 95-100; 125bp; English.
XX
CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
CC converting enzyme is a zinc metalloproteinase that plays roles in blood
CC pressure regulation and fertility. Zace2 can be expressed by standard
CC recombinant methodology. Zace2 polypeptides are useful for treating an
CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
CC diseases associated with inflammation like arthritis and enterocolitis,
CC as targets for identifying new angiotensin-converting enzyme (ACE)
CC inhibitors, and as a basis for rational drug design for inhibitory
CC molecules. The nucleic acids can be used to detect the expression of a
CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
CC for detecting and localizing Zace2 gene expression in tissue samples,
CC to determine whether a subject's chromosomes contain a mutation in the
CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
CC Inhibitors of ACE are used for treating hypertension of various
CC conditions, including left ventricular systolic dysfunction, progressive
CC renal impairment, scleroderma renal crisis, congestive heart failure due
CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
CC used to treat infertility while Zace2 antagonists are used for inducing
CC infertility. The present sequence represents a cDNA encoding the human
CC Zace2 protein.
XX
SQ Sequence 3334 BP; 1011 A; 640 C; 754 G; 929 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 3334
XX Score: 3869.00 Matches: 720
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-978-385-2_COPY_19_738 (1-720) x AAC84366 (1-3334)
XX
QY 1 SerPrrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 89 TTCACCATGAGAGAACAGGCCAAGACATTTTGGACAAATTTAAACCAAGAGCCGAAGAC 148
QY 21 LeuPheTyrrGlnSerSerIleuAlaSerTyrAsnTyrrAsnThrAsnIleThrGluLysAsn 40
Db 149 CTGTTCTATCAAAAGTCACTGCTGCTTGGAAATTAATACCAATATTACTGAAGAGAAAT 208
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db 209 GTCCAAAACATGATTAATGCTGGGGAGACAAATGCTGCTCTTTTAAAGAAACAGTCCACA 268
QY 61 LeuAlaGlnMetTyrrProLeuGlnGluIleGlnAsnLeuPrrValLysLeuGlnLeuGln 80
Db 269 CTTGCCCAATATGTTCCACTACAGAAATTCAGATCTCACAGTCAAGCTTCACTGCTCAG 328
QY 81 AlaLeuGlnLysAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db 329 GCTCTTCAGCAAAATGGGTCTTCAGTGTCTCAGACAGCAAGCAAGCAAGCTTGAACACA 388
QY 101 IleLeuAsnThrMetSerThrIleTyrrSerThrGlyLysValLysAsnProAspAsnPro 120
Db 389 ATTCTAAATACAAATGAGACCATCATCTACAGTATGGAAGAGTTGTATACCCAGATTAATCCA 448
QY 121 GlnGluCysLeuLeuLeuGluProGluLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db 449 CAAAGATCTTATTACTTGAACAGGTTTGATATTAATGCAAAACAGTTTGAAGTAC 508
QY 141 AsnGluArgLeuTrpAlaTrpGluSerTyrArgSerGluValGlyLysGlnLeuArgPro 160
Db 509 AATGAGAGAGCTCTGCGGCTTGGGAAAGCTGGAGATCTGAGGTGCGCAAGAGCTGAGGCCA 568
QY 161 LeuTyrrGluGluTyrrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrrGluAsp 180

```

Db	569	TTATATGAAGAGTATGCTGCTTGGAAAAATGAGAGCGCAAGACGAAATCATATTATGAGAC	628
OY	181	TyrGlyAspArgTyrTrpArgGlyAspArgGluValAlaSerGluTyrAspArgTyrSer	200
Db	629	TATGGGGGATTTATGGAGAGAGACTATAGTAAGTAATGGGGTGAAGATGGCTATGACTACACG	688
OY	201	ArgGlyGlnLeuIleGluAspValGluIleuIstThrPheGluGluIleLeuProLeuTyrGlu	220
Db	689	CGGGCCAGTTGATGATGAAGATGTGGACATACCTTTGAAGAGATTAAACATTATATGAA	748
OY	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro	240
Db	749	CATCTTCATGCGCTATGATGAGGGCAAAAGTGTGAATGGCTTACCTTCCTATATACGTCCA	808
OY	241	IleGlyCysLeuProAlaHisLeuLeuGluYAspMetCTrpGlyArgPheThrTrpAsnLeu	260
Db	809	ATTGGATGCCCTCCCGCTCATTTGGTTGGTATGATGTGGGGTGAAGATTTTGGCAAAATCG	868
OY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	869	TACTCTTTGACAGTCCCTTTGGACAGAAACCAACATAGATGTACATGATGCATATGGTGTG	928
OY	281	AspGlnAlaTrpAspAlaGlnArgIlePheLysGluValGluLysPhePheValSerVal	300
Db	929	GACCAAGCCCTGGGATGCACAGAGATATTACAGAGAGGCCGGAAGATCTTGTGTATCTGTT	988
OY	301	GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320
Db	989	GGCTCTTCCTAATATGACTACAGAGATTTGGGAAATTCATAGTCTAACGACCAGCAAAAT	1048
OY	321	ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTle	340
Db	1049	GTTCAGAAAGAGTCTGGCATCCCAACACTTGGGACCTGGGGAGAGGGGAGACTTCAGGATC	1108
OY	341	LeuMetCysTrpLysValThrMetCysAspPheLeuThrAlaHisHisGluMetGlyHis	360
Db	1109	CTTATGTGTACAAAGGTGACAAATGACAGACACTTCCTGACAGCTCATCATGAGATGGCGAT	1168
OY	361	IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAlaSerGlu	380
Db	1169	ATCCAGTGTGATATGGCATATGCTGTGCACACCTTTTCTGTCTAAGAAATGGAGCTAATGAA	1228
OY	381	GlyPheHisGluAlaValGluIleLeuSerLeuSerAlaAlaThrProLysHisLeu	400
Db	1229	GGATTTCCATGAAGCTGTGGGGAATCATGTCACTTCTGTGCAGCCACACTTAAGCATTTTA	1288
OY	401	LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe	420
Db	1289	AAATTCATATGGCTCTTGTGCACCCGATTTTCAAGAAAGCAATGAAACAGAAATTAACCTTC	1348
OY	421	LeuLeuLysGlnAlaLeuThrIleValGluIleThrLeuProPheThrTyrMetIleGlnLys	440
Db	1349	CTGGCTCAACACAGCACTCACATATGTTGGGACTGTGCCATTTACTTACATGTGTTAAGAAAG	1408
OY	441	TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp	460
Db	1409	TGGAGGTGAGTGTCTTTAAAGGGGAATTCOCACAAAGACAGTGAATGAAAAAGTGGTGG	1468
OY	461	GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys	480
Db	1469	GAGATGAAGCAGAGAGATGTTGGGTGTGGAACTGTGCCCAAGATGAAGAAACATACGTGT	1528
OY	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg	500
Db	1529	GACCCCGCATCTGTGTCCATGTTCTATATGATTACTATTCATTCGATATTCACAAAGG	1588
OY	501	ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGluPro	520
Db	1589	ACCCCTTATCCATTTCCATTTCAAGAAACACTTTGTCAAGACAGTAAACATGAAGCCCTT	1648
OY	521	LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu	540

Db	1649	CTGCACAATGTGACATCTCAACGCTCACGAACACTGTGACAGAAACTGTTCAATATCTGTG	1708
QY	541	ArgLeuGlyLysSerGluProThrPheLeuAlaLeuGluAsnValValGlyAlaIleAsn	560
Db	1709	AGGCTGGAAATCATCGAACCCCTGGACCTACCATTTGGAAATGTTGTATGGAGCAAGAAC	1768
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLeuLysAspGln	580
Db	1769	ATGATATGTAAAGCCACCTGCTCAACTACTTGTAGCCCTTATTACTTGCTGAAAGACACAG	1828
QY	581	AsnLysAsnSerPheValGlyTyrPserThrAspTrpSerProTyrAlaAspGlnSerIle	600
Db	1829	AACCAAGAAATCTTTTGTGGATGGAGATGCCAGCTGGAGTCCATATGCAAGACCAAGCATC	1888
QY	601	LysValArgIleSerIleuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn	620
Db	1889	AAAGTGAGATTAAGCCTAAATCAAGCTCTTGAGATTAAGCATTAATGAATGAGAACAGCAAT	1948
QY	621	GluMetTyrLeuPheAspArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640
Db	1949	GAAATGTATCCCTGTTCGGATCGATCGTTCGCAATGCTATAGAGCGAGTACTTCTTTTAAAGTA	2008
QY	641	LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg	660
Db	2009	AAAAATCGATGATTTCTTTTGGGAGAGAGATGTGCCAGTGGCTAATTTTAAACCAAGA	2068
QY	661	IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr	680
Db	2069	ATCTCCTTAATTTCTTGTGTGCTACGACCTAAATATGTCGTGATATCATCTCTAGAACT	2128
QY	681	GluValGlyLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700
Db	2129	GAGTGTGAAAGGCCATCAGGATGTCCCGGAGCCGTATCAATGATGCTTCCGTCCTGAAAT	2188
QY	701	AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro	720
Db	2189	GACAACACGCTAGAGTTCTTGCGGATACAGCCAAACACTTGACCTCTCAACCAACCCCTT	2248
RESULT 3			
AA12764			
ID	AA12764	standard; cDNA; 3396 BP.	
XX	AA12764;		
AC			
XX	25-JUL-2000	(first entry)	
DT			
XX			
DE	cDNA encoding a human angiotensin converting enzyme-2 (ACE-2).		
XX			
KW	Human; angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang. (1-9);		
KW	blood pressure; hypertension; congestive heart failure; atherosclerosis;		
KW	chronic heart failure; acute heart failure; myocardial infarction;		
KW	renal failure; SS.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	82..2499	
FT		/*tag= a	
FT		/product= "angiotensin converting enzyme-2"	
FT	sig_peptide	82..135	
FT		/*tag= b	
XX			
PN	WO200018899-A2.		
XX			
PD	06-APR-2000.		
XX			
PF	29-SEP-1999;	99WO-US22976.	
XX			
PR	30-SEP-1998;	98US-0163648.	
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Action LS, Robison KE, Hsieh FY;		

QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln 580
 DB 1816 ATGAATGTAAGCCAGCTCACTACTTGGAGCCCTATTACTGCGCAAGAACAGCAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAGAAATCTCTTTGGGATGGAGTACGACACGAGTGGATGCTCATATGCAACCAAGCATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysValTyrGluTrpAsnAspAsn 620
 DB 1936 AAATGAGGATTAACCTTAATATGCTCTTGGAGATTAAGCATATGATGAAAGACAAAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACCTGCTCCGATCATCTGTGCAATGCTATGAGGAGGAGTCTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATCAGATGATCTTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTAAATTTCTTGTGCACGACCTTAATATGCTGATATCATCTTAGAAT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAAGTGAAGAGCCATCAGAGATGCTCCGAGCCGATCATATGATGCTTCCGTCGAAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACACAGCCTTAAGTTCTGGGATACAGCCAAACCTTGACCTCTTAACGAGCCCT 2295
 RESULT 4
 AAD02758 standard; cDNA; 3396 BP.
 ID AAD02758;
 AC AAD02758;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Human angiotensin converting enzyme-2 (ACE-2) cDNA.
 XX
 KW Human; angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;
 KW screening; therapy; hypertension; congestive heart failure; CHF;
 KW inflammation; pain; ss.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT CDS
 FT 82..2499
 FT /tag- a
 FT /product- "Human angiotensin converting enzyme-2
 FT (ACE-2)"
 FT /EC_number- "3.4.15.1"
 FT /note- "this region is referred as SEQ. ID. NO.3 and is
 FT specifically claimed in claim 26"
 FT sig-peptide
 FT 82..135
 FT /tag- b
 FT mat-peptide
 FT 136..2496
 FT /tag- c
 FT /product- "Human mature angiotensin converting enzyme-2
 FT (ACE-2)"
 XX
 PN US6194556-B1.
 XX
 PD 27-FEB-2001.
 XX
 PF 11-DEC-1997; 97US-0989299.
 XX
 PR 11-DEC-1997; 97US-0989299.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.

XX
 PI Acton SL, Robison KE.
 XX
 DR WPI, 2001-210604/21.
 DR P-SDB; AAY72667.
 XX
 PT Novel genes encoding angiotensin converting enzyme-2 useful as
 PT antiseize or antigenic agents for therapeutics, diagnostics and
 PT screening assays.
 PS Claim 1; Fig 1; 76pp; English.
 XX
 CC The present sequence is human angiotensin converting enzyme-2 (ACE-2)
 CC cDNA. ACE is also referred as peptidyl dipeptidase A. Nucleic acid
 CC sequence encoding ACE-2 is useful as antiseize or antigenic agents for
 CC sequence specific modulation of gene expression or in the analysis of
 CC single base-pair mutations in the gene. Nucleic acid sequence encoding
 CC ACE-2 is useful in therapeutics, diagnostics and in screening assays.
 CC ACE-2 antagonist is used to treat hypertension or congestive heart
 CC failure (CHF). ACE agonist is used to reduce the inflammation and pain
 CC resulting from an insect sting or bite, which was accompanied by an
 CC injection of bradykinin. Anti ACE-2 antibodies are used to monitor ACE-2
 CC protein levels for determining the disease or condition associated with
 CC an aberrant protein level.
 XX
 SQ Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAD02758 (1-3396)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 136 TCCACCAATGAGAGACAGCCAGCAATTTTGGACAGTTTAACCAAGAGCCGAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 196 CTGTTCTTCAAAAGTCTCACTGCTTCTTGGATTAATACCAATATATAGAGAGAAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 DB 256 GTCCAAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGGAAAGCTCCACA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 316 CTGGCCCAATGATATCCACTACAGAAATTCAGATGTCACAGTCACAGCTTCCACTGAC 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCTCTTCAGCAAAAGGCTCTTCAGTCTGTCACAAAGACAGCAAAACGGTTGAAACACA 435
 QY 101 IleLeuAsnTrpMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 DB 436 ATCTAATATACAAAGAGACCACTACAGTCTGGAAGAAAGTTTAAACCAAGATATCCA 495
 QY 121 GlnGluLysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 496 CAAGATGCTTATTAATCTTGAACCAAGGTTGAATGAATATGCAAAACAGTTTAAGCTAC 555
 QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 DB 556 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTAGGTCGCAAGAGCTGAGGCCA 615
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTATATGAAGATATGCTGCTTGAAGAAATGAGATGCAAGAGCAATATATATGAGGAC 675

QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 676 TANGGGATATTATGGAGAGAGACTAATGAAGTAAATGGGTAGATGGCTATGACTACAC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluIleHisThrPheGluGluIleLeuProLeuTyrGlu 220
 DB 736 CGGGCCAGTGTGATTGAAGATGGGAACATACCTTTGAAGAGATTAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyrValArgAlaLeuLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCCATATGGAGAGCAAAAGTTGATGAAATCCCTATCCCTATATCATCTCA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260
 DB 856 ATTGAGATGCCCTCCCTGCTCATTTGCTGGATATATGGGGTGGATTTGGACAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnIleProAsnIleAspValThrAspAlaMetVal 280
 DB 916 TACTCTTGGACATGCCCTTTGGACAGAAACCAACATAGATGTCTGATGCAATGGTG 975
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLeuGluAlaGluLeuPhePheValSerVal 300
 DB 976 GACACAGCCCTGGAGATGACAGACAAATATTCAGAGAGCCGAGAAAGTTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGCTCTTCCATATGACTCAAGATCTCGGAAATATCCATGCTAACGAGCCAGCAAAAT 1095
 QY 321 ValGlnIleValAlaValCysHisProThrAlaTrpAspLeuGlyGlyAspPheArgIle 340
 DB 1096 GTTCAGAAACAGCTGCTCCATCCACAGCTTGGAGACTGGGAAAGGCGCTTCAGAGATC 1155
 QY 341 LeuMetCysThrIleValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHis 360
 DB 1156 CTATATGCGCAAAAGGAGCAATGGAGAGACTTCTGACACCTCATCATGATGAGTGGGCAT 1215
 QY 361 IleGlnIleThrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1216 ATCCAGATATGATGGCATATGCTGCGCAACCTTTCTGCTAAGAAATGAGCATATGAA 1275
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCATGAAAGCTGTGGGGAATCATCTCTCTGACAGCACCTTAAGCATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGluThrGluIleAsnPhe 420
 DB 1336 AAATCCATGTGGTCTGTCACCGCATTTTCAAGAAGACATGAACAGAAATAAACTTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAAAACAGACACACAGATGTTGGAGACTCTGCCATTTACTTACATGTTAGAGAAC 1455
 QY 441 TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp 460
 DB 1456 TGGAGGTGGATGCTTTTAAAGGGAAATTCCTCAAAACAGAGTGGATGAAAGAGTGTGG 1515
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAAGCAGAGATAGTTGGGGTGGTGAACCTGCCCCATGATGAACATACTGT 1575
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArg 500
 DB 1576 GACCCCGCATCTGCTTCCATGTTCTAATGATTACATTCATTCATATTTACACAGG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCATATCCAGTTTCAAGAAGCACTTTGTACAGCAGCTAAACATGAAGGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTCGACAAATGTGACATCTCAAACTACAGAAAGCTGACAGAAACCTGTTCAATATCTGT 1755
 QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsn 560

DB 1756 AGGCTTGAAATATCGAACCCCTGACCCCTACATTTGAAATGTTGAGGCAAAAGAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln 580
 DB 1816 ATGAATGTAAGCCCTGCTGACACTTGTAGCCCTTATTTACTGGCTGTAAGACACAG 1875
 QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAGAAATTTCTTTGTGGATGAGTACCCGACTGAGTCCATCATATGACAGCAAAAGCATC 1935
 QY 601 LysValAlaGlyIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluIleTrpAsnAspAsn 620
 DB 1936 AAAGTGAGATTAAGCTTAAATACACCTTTGGAGTAAAGCATTAAGATGAGACGACAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACTCTTCCGATCATCTGTGATATGCTATGAGAGCACTTTTAAAGTA 2055
 QY 641 LysAsnGluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATGAGATGATTTCTTTTGGAGAGAGATGTCGAGTGAATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCCTTAAATTTCTTTGTCTACTGACCTAAATATGCTGATATCATCTTCAAGACT 2175
 QY 681 GluValGlyLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAACTTGAAAGGCCATCGATCGATGTCGCCGAGCCCTATCATGATGCTTCGCTGAT 2235
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGluProPro 720
 DB 2236 GACAACAGCGCTAGAGTTCTGGGATACAGCCACACTTGACCTCTTAACACAGCCCT 2295
 DB 2295 GACAACAGCGCTAGAGTTCTGGGATACAGCCACACTTGACCTCTTAACACAGCCCT 2295
 RESULT 5
 ABK87623
 ID ABK87623 standard; cDNA: 3396 BP.
 AC ABK87623;
 XX XX
 DT 24-SEP-2002 (first entry)
 XX XX
 DE cDNA encoding human angiotensin converting enzyme-2 (ACE-2) protein.
 KW Human; angiotensin converting enzyme-2; ACE-2; body weight disorder;
 KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;
 KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;
 KW familial partial lipodystrophy; hypercholesterolaemia; hyperlipidaemia;
 KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;
 KW neurodegenerative disorder; peptide hormone; cytokine processing;
 KW myocardial infarction; cardiovascular pathology; inflammatory bowel disease;
 KW systemic inflammation response syndrome; polytrauma; pain; stroke;
 KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;
 KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;
 KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;
 KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;
 KW adult respiratory distress syndrome; wound healing; appetite;
 KW body mass index; gene; ss.
 XX XX
 OS Homo sapiens.
 FH
 FH key location/Qualifiers
 FT CDS 82..2499
 FT /tag- a
 FT /product- "human angiotensin converting enzyme-2 (ACE-2)
 FT sig-peptide 82..138
 FT /tag- b
 FT mat-peptide 139..2496
 FT /tag- c
 XX XX
 PN W0200239997-A2.

XX 23-MAY-2002.
 XX 31-OCT-2001; 2001WO-US45703.
 XX 01-NOV-2000; 2000US-0704216.
 XX 29-MAY-2001; 2001US-0870382.
 XX 19-OCT-2001; 2001US-371741P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Acton ST, Ocala FL, Gould AE, Dales NA, Guan B, Brown JA;
 XX Patane M, Kadambi VU, Solomon M, Stricker-Krongrad A;
 XX WPI: 2002-547572/58.
 XX P-PSDB: AAU99701.
 XX Treating body weight disorder and increasing muscle mass comprises
 XX administering angiotensin converting enzyme-2 modulating compound
 XX Disclosure: Page 383-387; 395pp; English.
 XX The present invention describes a new method of treating a body weight
 XX disorder, increasing muscle mass and decreasing body fat by
 XX administration of angiotensin converting enzyme (ACE)-2 modulating
 XX compound. The invention can be used for treating body weight disorders,
 XX particularly obesity of at least grade I, diabetes, atherosclerosis and
 XX a state associated with lipid metabolism. The method is used for treating
 XX rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,
 XX generalised partial lipodystrophy, familial partial lipodystrophy,
 XX hypercholesterolaemia, hyperlipidaemia, an aberrant metabolic rate,
 XX congestive heart failure, chronic heart failure, left ventricular
 XX hypertrophy, acute heart failure, neurodegenerative disorders (e.g.
 XX Alzheimer's disease, Parkinson's disease and Huntington's disease),
 XX diseases associated with peptide hormones or cytokine processing,
 XX myocardial infarction, cardiomyopathy, systemic inflammation response
 XX syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and
 XX chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis
 XX and periodontal disease, dysmenorrhoea, premature labour, brain oedema
 XX following focal injury, diffuse axonal injury, stroke, reperfusion
 XX injury, cerebral vasospasm after subarachnoid haemorrhage, allergic
 XX disorders including asthma, adult respiratory distress syndrome, wound
 XX healing and scar formation. The invention decreases the appetite,
 XX increases muscle mass and decreases body fat of subject having body mass
 XX index of greater than 23 (preferably 24.9)kg/m². The present nucleic
 XX acid sequence encodes the human ACE-2 protein of the invention.
 XX
 XX Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 0 Length: 3396
 XX Score: 3869.00 Matches: 720
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 100.00% Indels: 0
 XX DB: 24 Gaps: 0
 XX
 XX US-09-978-385-2_COPY_19_738 (1-720) x ABR67623 (1-3396)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 136 TCCACCACTGAGCAACGCGCAAGACATTTTGGACAAAGTTTAACCGACGAGCGGAAGAC 195
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnHisIleThrGluGluAsn 40
 DB 196 CTGCTCTATCAAAAGTTCACCTGCTTCGTAATTAACACCAATATTAAGTGAAGCAAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 256 GTCCAAACATGAAATATGCTGGGAGCAAAATGCTGCTTTTAAAGGACGTCACCA 315
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

DB 316 CTTCGCCAAATGATCACTACAAAGAAATTCAGAAATCTCAAGCTTCAAGCTTCAGCTGAC 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 376 GCTCTTACGAAATATGGCTTCAGCTGCTGCAAGAGCAAGCAAGCAAGCTTGAACACAA 435
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 436 ATTCTAATATCAATGAGCACCACATCTACAGTACGAGAAAGTTTGTAACCCAGATTAATCCA 495
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 496 CAAGAAATGCTTATTAAGTGAACAGCTTGAATGAATGAATGAATGAATGAATGAATGAAT 555
 QY 141 AsnGluArgLeuThrPheAlaTyrPheGluSerTyrPheSerGluValGlySerGluLeuArgPro 160
 DB 556 AATGAGAGGCTGCTGGCTGGGAAACCTGAGATCTGAGCTGGCAAGACGTCGAGCCCA 615
 QY 161 LeuTyrGluGluTyrValValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 616 TTATATGAGAGATATGCTGCTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 675
 QY 181 TyrGlyAspTyrTyrPheGlyAspTyrGluValAsnGlyValAspLysTyrAspTyrSer 200
 DB 676 TATGGGATTAATGAGAGAGAGACTATGAAGTAAAGGGGTAGATGCTATGACTATACAC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 736 CGGGCCAGTATTAATGAATGAGACATGAGACATGAGACATGAGACATGAGACATGAGAC 795
 QY 221 HisLeuHisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 796 CATCTTCATGCTATGAGAGGCAAGTGAATGAATGCTATGCTATGCTATGCTATGCTATG 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 DB 856 ATGGATGCTCTCTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 915
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 916 TACTCTTACAGTTCCTTGGACAGAAACCAACATGATGTTACTGATGATGATGATGATG 975
 QY 281 AspGlnAlaTyrPheAspAlaGlnArgIlePheLysGluAlaGluLysPheValSerVal 300
 DB 976 GACCAAGCCTGGAGTACAGAGAAATATCAAGAGAGCGCAAGACTCTTGATATCTGT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 1036 GGTCTTCTTAATATGACTCAAGAGATTCGGGAAATTCATCCTTAACGACCGCCAGAAAT 1095
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheAspLeuGlyLysGlyAspPheArgTle 340
 DB 1096 GTTCAGAAAGAGCTGCTGCATCCACAGCTTGGACCTGGGAGAGGCGATTCACGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetCylHis 360
 DB 1156 CTTATGTGCAAAAGTGCATATGAGACACTTCGACAGCTCATCATGATGAGGCGCAT 1215
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuAsnGlyValAsnGlu 380
 DB 1216 ATCCAGATATGATATGATGATATGCTGACATCTTTCGCTTAAGAAATGGAGCTTAATCAA 1275
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1276 GGATTCATGAAAGCTGTTGGGAAATCATGTCATCTTTCGACGACACACCTTAACATTTA 1335
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1336 AATCCATGCTGCTCTCTCAACCCGATTTTCAAGAGCAATGAAGAGAAATTAATCTTC 1395
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1396 CTGCTCAACCAAGCCTACGATGCTGGAGCTGCGCATTTACTTACATGTTAGAGAG 1455

QY 441 TTPATGTPMetValPheLysGlyLueProLysAspGlnTTPMetLysTTPTrp 460
 DB 1456 TGGAGGTGAGTCTTTAAAGGGGCAATTCACCAAGACAGTGGATGAAAAAGTGGG 1515
 QY 461 GUMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
 DB 1516 GAGATGAGGAGGAGATGTTGGGGTGGAGACCTGGCCCATGATGAAACATATCTGT 1575
 QY 481 ASPProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 DB 1576 GACCCCGCATCTCTTCATCTGTTCTCATGATTAATCATCATCATCATATACCAAG 1635
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeucysGlnAlaAlaLysHisGluGlyPro 520
 DB 1636 ACCCTTACCAATTCACATTCATTCACAGAACCTTTGTCAGACGACCAATGAAAGGCCCT 1695
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
 DB 1696 CTGCACAAATGTGACATCTCAAACTCTACAGAACCTGACAGAACTGTTCAATATCTGT 1755
 QY 541 ArgLeuGlyLysSerGluProTyrPheLeuAlaLeuGluLysValValGlyAlaLysAsn 560
 DB 1756 AGGCTTGAATAATCAGAACCTGGACCTTACCATTTGAAATGTTTGTAGAGCAAAATAC 1815
 QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGln 580
 DB 1816 ATGAAATGTAAGCCAGCTGCTCACTACTTGAAGCCCTTATTTACCTGGTGAAGAGCCAG 1875
 QY 581 AsnLysAsnSerPheValGlyTyrPheSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
 DB 1876 AACAAAGATTTCTTTTGGGAGTGAAGTACGAGTGCATATGACAGCAAGCATC 1935
 QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAspAsn 620
 DB 1936 AAGGTGAGATTAAGCCCTTAAATCAGCTTGGAGATTAACATATGAAATGAAAGCAAT 1995
 QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
 DB 1996 GAAATGTACTGTTCCGATCATCTGTTGCATATGCTATGAGGAGTACTTTTAAAGTA 2055
 QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
 DB 2056 AAAAATCAGATGATCTTTTGGGAGGAGATGTCGAGTGCATATTTGAAACCAAGA 2115
 QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
 DB 2116 ATCTCTTAATTTCTTTGTCACTGCACTTAAATAATGTCTGATATCTATCTAGAACT 2175
 QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
 DB 2176 GAAGTTGAAGAGCCATCAGAGATGTCGAGAGCCGATCATATGATCTTTCCGCTGAAAT 2235
 QY 701 AspAsnSerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2236 GACACAGACCTTACAGATTTCTGGGATACAGCCAACTTGAACCTCTTAAACAGCCCT 2295

RESULT 6
 AAD32586
 ID AAD32586 standard; cDNA; 3396 BP.
 AC AAD32586;
 XX 18-JUN-2002 (first entry)
 DE Human ACE-2 full-length cDNA.
 XX
 KW Human: angiotensin converting enzyme-2; ACE-2; therapy; hypertension;
 KW peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;
 KW myocardial infarction; heart failure; arrhythmia; renal failure; gene;
 KW inflammation; fertility; enzyme; X chromosome p21-22; ss.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 5'UTR 1..81
 FT CDS /tag- a
 FT 82..249 /tag- b
 FT /product= "Human ACE-2 protein"
 FT 82..135
 FT mat_peptide /tag- c
 FT 135..249 /tag- d
 FT /product= "Human mature ACE-2 protein"
 FT /EC number= "3.4.15.1"
 FT 136..2301 /tag- e
 FT /note= "Encodes extracellular domain"
 FT 1201..1215 /tag- f
 FT /note= "Encodes zinc binding domain (ZBD)"
 FT 2302..2376 /tag- g
 FT /note= "Encodes transmembrane domain"
 FT /tag- h
 FT /note= "Single chain polymorphism: This polymorphism results in an amino acid change from asn to asp"
 FT 2377..2496 /tag- i
 FT /note= "Encodes cytoplasmic domain"
 FT 2500..3396 /tag- j
 FT replace (2834, T) /tag- k
 FT /note= "Single chain polymorphism"
 PN WO200212471-A2.
 PD 14-FEB-2002.
 XX 09-AUG-2001; 2001WO-US25059.
 XX 09-AUG-2000; 2000US-0635501.
 PR (MILL-) MILLENNIUM PHARM INC.
 PA Action S, Robison KE, Hsieh FY;
 PI WPI: 2002-257481/30.
 DR P-PDB: AAE20353.
 XX
 XX Isolated human polypeptide, known as angiotensin converting enzyme-2,
 PT useful for treating or preventing the development of an abnormal blood
 PT pressure or related diseases, e.g. hypertension, heart failure or
 PT myocardial infarction
 PS Claim 1; Fig 1; 218pp; English.
 XX
 XX The invention relates to human angiotensin converting enzyme-2 (ACE-2)
 CC polypeptides and polynucleotides. ACE-2 is also known as peptidyl
 CC dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
 CC for treating or preventing the development of abnormal blood pressure
 CC and diseases or disorders associated with the protein in a subject. The
 CC diseases include hypertension, hypotension, congestive heart failure,
 CC atherosclerosis, arrhythmia and renal failure. They are also useful
 CC for treating inflammatory conditions and diseases relating to fertility.
 CC The present sequence is human full-length ACE-2 cDNA. The ACE-2 gene is
 CC located on the X chromosome at p21-22.
 SQ Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
 Alignment Scores: 0 Length: 3396
 Pred. No.:

Score: 3669.00
Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 100.00%

Matches: 720
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AMD32586 (1-3396)

QY 1 SerThrIleGluGluGlnAlaIleValThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
DB 136 TCCACCATTCAGAGAACAGCCAGACATTTTGGACAAAGTTTAAACCCAGAGCCGAGAC 195
QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsn 40
DB 196 CTCTTCTATCAAGGTTCACTTCTCTTCTGGAATTAACACCAATTTACTGAGAGAAAT 255
QY 41 ValGlnAsnMetAsnAsnAlaGluAspLysTyrPserAlaPheLeuLysGluGlnSerThr 60
DB 256 GTCCAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 315
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB 316 CTTCGCCAAATGATCCACTACAGAAATTCAGAAATTCACACTCAAGCTTCAGCTCAG 375
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
DB 376 GCCTTCAGCAAAATGGCTCTCAGTGCCTGTCAGAAAGACAAAGCAAAACGTTGAGACACA 435
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
DB 436 ATCTTAATATCAATGACACCACTACAGTACAGTGAAGTTGTAACCCAGATTAATCCA 495
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
DB 496 CAGAAATGCTTATTAATCTTGAACCAAGTTTGAATGAATATGCAAAACAGTTTAGACTAC 555
QY 141 AsnGluArgLeuThrPalatrpGluSerTyrPargSerGluValGlyLysGlnLeuAspPro 160
DB 556 AATGAGAGGCTCTGCGCTGGGAAAGCTGGAAGTACAGTACGCTGCGCAAGCAGCTGAGGCCA 615
QY 161 LeuTyrGluGluTyrValIleValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
DB 616 TTATATGAGAGATGTTGGTCTTAAATAATGATGCGAAGCAAGCAATCTATTGAGGAC 675
QY 181 TyrGluAspTyrTyrPargGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
DB 676 ATATGGGATATATGAGAGAGACTATGAAGTAATGCGGATGCTATGACTACAGC 735
QY 201 ArgGluGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
DB 736 CGCGGCCAGTTGATGGAAGATGTGAACATACCTTTGAAGAGATTAACCACTATATGAA 795
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
DB 796 CATCTTCATGCTATGTCAGGCAAAAGTGTATGATGCTATCTCTCTATATACAGTCCA 855
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrPglYargPheTyrPheAsnLeu 260
DB 856 ATTGATGCTCTGCTGCTCATTTGCTGCTGATATGCGGGTAAATTTTGGACAAATCTG 915
QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
DB 916 TACTCTTTGACAGTTCCTTTGGACAGAAACCAACATAGATGTTACTGATGCAATGCTG 975
QY 281 AspGlnAlaTyrAspAlaGlnArgTlePheLysGluAlaGluLysPhePheValSerVal 300
DB 976 GACCAAGGCTGGGATGACAGAGAAATATTCAGAGGCGGAGAAAGTCTTGTATCTGTT 1035
QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPglLysAsnSerMetLeuThrAspProGlyAsn 320
DB 1036 GCTCTTCTTAATATATACATCAAGGATTCGGGAAATTCATGCTACAGCAGCCAGGAAT 1095
QY 321 ValGlnLysAlaValLysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgTle 340

DB 1096 GTTCAGAAAGCAGCTGCGCATCCACAGCTTGGGACCTGGGAGGCGGCTTCAGAGATC 1155
QY 341 LeuMetCysThrLysValIleThrMetAspPheLeuThrAlaHisHisGluMetGlyHis 360
DB 1156 CTATATGTCACAAAGGTACATATGAGACACTCTCTGACAGCTATCATGATGAGGGCAT 1215
QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
DB 1216 ATCCAGTATGATATGGCATATGTCACAAACCTTTCTGTAAGAAATGAGAGCTAAATGAA 1275
QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerIleAlaIleProLysHisLeu 400
DB 1276 GGATTCATGAAAGCTGTGGGAAATCATGCTCTTCTGACGCCACCACTTAACATTTA 1335
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLeuAsnPhe 420
DB 1336 AATCATATGGCTCTCTGCTACCGCATTTTCAAGAGACATGAAACAGAAATTAACCTTC 1395
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
DB 1396 CTGCTCAAAACAGCACATCAGATTTGTTGGACTGCTGCTATTAATGTTAGAGAAC 1455
QY 441 TrrArgTyrMetValPheLysGlyGluIleProLysAspGluTrrPheLysTyrTrrP 460
DB 1456 TGGAGGTGATGCTTTTAAAGGGAATTTCCCAAGACCGAGTGATGAAAAAGTGGTGG 1515
QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
DB 1516 GAGATGAAAGCGAGATATGTTGGGTGGTGAACCTGCGCCATGATGAACATATACGT 1575
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
DB 1576 GACCCCGCATCTCTGCTCATGTTCTTAATGATTAATCATTCATTCGATTAACAGAG 1635
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
DB 1636 ACCCTTACCAATTCAGATTCAAGAGACATTTGCAAGCAGCTTAACATGAAGGCCCT 1695
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
DB 1696 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGTGGACAAACCTGTCAATGTGTG 1755
QY 541 ArgLeuGlyLysSerGluProTrrPheLeuAlaLeuGluAsnValIleGlyAlaLysAsn 560
DB 1756 AGGCTTGGAATACAGAACCTTCGACCTCAGACCTTGGAAATGCTGTGAGAGCAAGAAC 1815
QY 561 MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrPheLysAspGln 580
DB 1816 ATGAATGTAAAGGCACCTGCTCAACTTTGAGCCCTTATTAACCTGCTGTAAGACCG 1875
QY 581 AsnLysAsnSerPheValGlyTrrPserTrrAspTrrPserProTyrAlaAspGlnSerIle 600
DB 1876 AACAAATCTTTTGTGGATGAGTACCGATGAGTCCATGACCAACCAAGACATTC 1935
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrPheAsnAspAsn 620
DB 1936 AAAGTGAAGATTAACCTTAATACAGCTTGGAGATGAAGCAATGAAGTGAAGCAAT 1995
QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
DB 1996 GAATGTACCTGTTCCGATCATCTGTGATATGATATGAGCAGACACTTTTAAAGTA 2055
QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
DB 2056 AAAAATCAGATGATCTTTTGGGAGAGAGATGTCGAGTGGCTAATTTGAACCAACA 2115
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
DB 2116 ATGCTCTTAATTTCTTTGTCACCTGACCACTAAAAATGTGTGATATCATCTTGAAGCT 2175
QY 681 GluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700

Db 2176 GAAGTGAAGGCGATGAGATGCCGAGCCGATATCATGATGCTTCGCTGCAAT 2235
 QY 701 ASPASerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2236 GACAAAGCGCTTACGTTCTGGGAGATACAGCAACACTTGGACCTCTTAACAGCCGCCCT 2295
 RESULT 7
 AAS14880
 ID AAS14880 standard; cDNA; 2920 BP.
 AC AAS14880;
 XX
 DT 20-DEC-2001 (first entry)
 DE Human cDNA encoding novel human protein NHP #1.
 XX
 KW Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;
 KW antiarthritic; cytosolic; antiarteriosclerotic; vulnary;
 KW neuroprotective; nootropic; antiparkinsonian;
 KW anti-human immunodeficiency virus; antiallergic; vasotropic; cardiant;
 KW hypotensive; anorectic; antileptin; neuroleptic; anticonvulsant;
 KW antianemic; immunosuppressive; cerebroprotective; antimicrobial;
 KW immunomodulator; antibacterial; antiparasitic; thymimetic;
 KW gastrointestinal disorder; cardiovascular; vasoconstriction;
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;
 KW hyperproliferative disorder; pulmonary system disorder;
 KW central nervous system disorder; bone disorder;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW Huntington's disease; schizophrenia; mania; dementia; paranoia;
 KW panic disorder; learning disability; amyotrophic lateral sclerosis;
 KW psychosocial; autism; sleep disorder; immune system disorders;
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;
 KW multiple sclerosis; ischemic brain injury; stroke; infectious disease;
 KW diabetes mellitus; immunological disorder; asthma; AIDS;
 KW acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
 KW neural system disorder; respiratory disorder; olfactory disorder;
 KW wound healing; chromosome X.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 213..2348
 FT /tag="a
 FT /product="NHP #1"
 FT /transl_except- (pos:867..869,aa:Xaa)
 FT /transl_except- (pos:930..932,aa:Xaa)
 FT /transl_except- (pos:1707..1709,aa:Xaa)
 FT /note="Xaa= Any amino acid"
 XX
 PN WO2001/4896-A1.
 XX
 PD 11-OCT-2001.
 XX
 PE 02-APR-2001; 2001WO-US10542.
 XX
 PR 03-APR-2000; 2000US-194118P.
 XX
 PR 29-SEP-2000; 2000US-236384P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 DR WPI; 2001-626394/72.
 DR P-PSDB; AAO09092.
 XX
 PT New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders

PS Claim 1; Page 291-292; 318pp; English.
 XX
 CC The invention relates to novel human proteins (NHP) and the
 CC nucleic acids that encode them and antibodies raised against them.
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
 CC prognosis, prevention and/or treatment of diseases and/or disorders
 CC involving vasoconstriction, gastrointestinal disorders, cardiovascular
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
 CC cachexia, disorders of small intestine, disorders of reproductive system
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
 CC and other hyperproliferative disorders, disorders of pulmonary system,
 CC central nervous system disorders, bone disorders, neurodegenerative
 CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
 CC panic disorder, learning disabilities, amyotrophic lateral sclerosis,
 CC psychoses, autism, sleep disorders), immune system disorders (e.g.
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
 CC central nervous system disorders (e.g. multiple sclerosis, ischemic
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence encodes an NHP of the invention and
 CC is located on the X chromosome.
 CC
 SQ Sequence 2920 BP; 897 A; 568 C; 654 G; 788 T; 13 other;
 Alignment Scores:
 Pred. No.: 0 Length: 2920
 Score: 3833.00 Matches: 715
 Percent Similarity: 99.31% Conservative: 4
 Best Local Similarity: 99.17% Mismatches: 1
 Query Match: 99.07% Indels: 1
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAS14880 (1-2920)
 QY 1 SerThrIleGlnGlnGlnAlaIleThrPheLeuAspLysPheAsnHisGlnIleGlnAs 20
 Db 83 TCCACGATGAGGAGCAAGCCAGCAACATTTGGGAGCAAGTTTACCAACGAGCCGAGGA 142
 QY 20 PLeuPheTygIlnSerSerLeuAlaSerTPAsnTyAsnThAsnIleThrgIlnAs 40
 Db 143 CCTGTTCTATCAAGATTCACCTTCTCTGGAATTAATTAACCAATTTACTGAAGGAA 202
 QY 40 nValGlnAsnMetAsnAsnAlaGlyAspIysTrpSerAlaPheLeuLysGlnIlnSerTh 60
 Db 203 TGTCCAAAACATGATATATGCTGGGAGCAAAATGCTCTTTTAAAGAACAGTCCAC 262
 QY 60 rLeuAlaGlnMetTygProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnIleu 80
 Db 263 ACTTGCCCAATGATATCCTACACAGAAATTCAGAAATCTACAGTCAAGCTTCACTGCA 322
 QY 80 nAlaLeuGlnGlnAsnGlySerValLeuSerGlnAspLysSerLysArgLeuAsnTh 100
 Db 323 GGCCTCTTCAGCAAAATGGGCTTCTCAGTCTCAGAGACAAGAGCAAAAGGTTGAACAC 382
 QY 100 rIleLeuAsnThrMetSerThrIleTygSerThGlyLysValCysAsnProAsnAsnPr 120
 Db 383 AATTCCTAATATCAATGAGCACCATCTACAGTACGAGAAAGTTGTAAACCAATATATCC 442
 QY 120 oGlnGlnCysLeuLeuLeuGlnIleProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTy 140
 Db 443 ACAAGAAATGCTTATTAATTGACACAGGTTTGAATGAATATATGCAAAAGTTTACACTA 502
 QY 140 rAsnGlnArgLeuThrPalatirpGlnSerTPAsnSerGlnValGlyLysGlnIleuArgpr 160
 Db 503 CAATGAGAGGCTCTGGGCTTGGGAGAAAGCTGAGATCTGAGAGTGGCAACGACGTCAGGCC 562
 QY 160 oLeuTygGlnGlnIleTygValValLeuLysAsnGlnIleAlaArgAlaAsnHisTygGln 180

KW neural system disorder; respiratory disorder; olfactory disorder;
 KM wound healing.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 213.998
 FT /tag- a
 FT /product- "NHP #11"
 XX
 XX MO200174896-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-US10542.
 XX
 PR 03-APR-2000; 2000US-194118P.
 PR 29-SEP-2000; 2000US-236384P.
 XX
 PA (HOMA-) HUMAN GENOME SCI INC.
 XX
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
 PI Li Y, Dillon PJ;
 DR P-PSDB; AA09102.
 DR WPI: 2001-626394/72.
 XX
 PT New human proteins, useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cardiovascular
 PT disorders, autoimmune disorders and reproductive disorders
 XX
 PS Claim 1; Page 297-298; 318pp; English.
 XX
 CC The invention relates to novel human proteins (NHP) and the
 CC nucleic acids that encode them and antibodies raised against them.
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,
 CC prognosis, prevention and/or treatment of diseases and/or disorders
 CC involving vasoconstriction, gastrointestinal disorders, cardiovascular
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
 CC cachexia, disorders of small intestine, disorders of reproductive system
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours
 CC and other hyperproliferative disorders, disorders of pulmonary system,
 CC central nervous system disorders, bone disorders, neurodegenerative
 CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paranoia,
 CC panic disorder, learning disabilities, immune system disorders (e.g.
 CC psychoses, autism, sleep disorders), amyotrophic lateral sclerosis,
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
 CC central nervous system disorders (e.g. multiple sclerosis, ischemic
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system
 CC disorders, respiratory disorders, olfactory disorders and wound
 CC healing. The present sequence encodes an NHP of the invention.
 XX
 SQ Sequence 2911 BP; 896 A; 570 C; 655 G; 788 T; 2 other;

Alignment Scores:

Pred. No.: 0 Length: 2911
 Score: 3785.00 Matches: 715
 Percent Similarity: 99.31% Conservative: 4
 Best Local Similarity: 99.17% Mismatches: 4
 Query Match: 97.83% Indels: 3
 DB: 22 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AAS14890 (1-2911)

QY 1 SerThrIleGluGlnAlaIleuThrPheLeu-AspLysPheAsnHisGluAlaGluAs 20
 DB 83 TCACACATTGAGAGAACAGCCAAAGCATTTTGGGACAAATTAAACCAAGAGCGGAGA 142

QY 20 pIeupheTYrGlnSerSerleuAlaSerTrpAsnTrpAsnIleThrGluGluAs 40
 DB 143 CCTGTTCTATCAAAAGTTCACTGCTTGGATTTAAACACCAATTTATGGAAGAGAA 202
 QY 40 nValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnIleSerTh 60
 DB 203 TGTCAAAACATGATTAATGCTGGGGCAAAATGGTGTGCTTTTAAAGAAACATCCAC 262
 QY 60 rIleuAlaGlnMetTYrProLeuGlnIleIleGlnAsnLeuThrValLysLeuGlnLeuG 80
 DB 263 ACTTGCCCAAAATGATACCACTACCAAGAAATTCAGATTCACAGTCAAGCTTCACCTCA 322
 QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
 DB 323 GGCTCTTCAGCAAAATAGGGCTTCACAGTCTTCAGAGACAAAGCAAAAGGTTGAAAC 382
 QY 100 rIleLeuAsnThrMetSerThrIleTYrSerThrGlyLysValLysAsnProAspAsnPr 120
 DB 383 AATTCCTAAATACAAATGAGCACCACATCTACAGTACTGGAAGTTGTAAACCCAGATTAATCC 442
 QY 120 oGlnGlnCysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTY 140
 DB 443 ACAAGATGCTTATTAATCTTGAACCAAGTTTGAATGAATTAATGCAACAGTTTGAAGCTA 502
 QY 140 rAsnGlnArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPr 160
 DB 503 CAATGAGAGCGCTCTGGGGCTTGGGAAGCTGAGATCTGAGCTGGCAAGCAGCTGAGGCG 562
 QY 160 oLeuTYrGlnGluTYrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTYrGluAs 180
 DB 563 AATTATATGAAAGTATGTGCTTGAATAATGATGCAAGACCAATCATTTATGAGGA 622
 QY 180 pTYrGlnAspTYrTrpArgGlyAspTYrGlnValAsnGlyLysGlyTYrAspTYrSe 200
 DB 623 CTAATGGGATTTTGGAGAGAGACTATGAAGTAATGGGTGAGTGGCTATGACTACG 682
 QY 200 rArgGlnLeuLeuIleGlnAspValGlnHisThrPheGlnGluIleLysProLeuTYrG 220
 DB 683 CCGCGCGCAGATTGATTGAAGATGTGGAACATATCCTTGAAGAGATTAACCATTAATGA 742
 QY 220 uHisLeuHisAlaTYrValAlaValGlyLeuMetAsnAlaTYrProSerTYrIleSerPr 240
 DB 743 ACATTTATGCTATGATGAGGCCAAGATTGATGATGCCATTCCTCTATACATGATGCC 802
 QY 240 oIleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpAsnLe 260
 DB 803 AATGGAATGCTCCCTGCTCATTTGCTGGGATGATGGGTGATTTGGACAAATCT 862
 QY 260 uTYrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVa 280
 DB 863 GTACTCTTTGACAGTCCCTTGGACAGAAACCAACATATGATGCTGATGCAATGCT 922
 QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerVa 300
 DB 923 GGACGAGCGCTGGAGTGCACAGAAATTTTAAAGAGCGGAGAAAGTTCTT-GRATGTGT 981
 QY 300 lGlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAs 320
 DB 982 TGTCTTCCTTAATATGACTCAAGAGATTCGGGAATAATTCATGCTAACGACCCAGGAA 1041
 QY 320 nValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgI 340
 DB 1042 TGTTCAGAAACCAATGTCACATCCACAGCTTGGAGCCTGGGGGAGGCGCACTTCAGAGAT 1101
 QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyH 360
 DB 1102 CCTATATGTCACAAAGGTGACAAATGAGAGCTTCTATACAGCTCATATGATGATGAGGGGCA 1161
 QY 360 sIleGlnTYrAspMetAlaTYrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnG 380
 DB 1162 TATCCAGATGATATGCAATGCTGCAACACCTTTCTGCTAAGAAATGAGAGCTATATGA 1221
 QY 380 uGlyPheHisGlnLysAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLe 400

```

Db      1222 AGGATTCACGAGCTGTTGGGAAATCATGTCTCTTTCAGACGACACCTTAACACTTT 1281
Qy      400 uylsSerIleGlyLeuLeuSerProApbheGlnIuAspaNgIuThrGluIleAsnPh 420
Db      1282 AAAATCCATGGCTCTCTGTCAACCCGATTTTCAAGAACATGAACAGAAATTAACCTT 1341
Qy      420 eleuLeuLyGlnAlaLeuThrIleValGlyThrLeuProbheThyMetLeuGluY 440
Db      1342 CCGTCTCAACAGCACTCAGATTTGTGGAGCTGCGCATTTACTTACATGTTAGACAA 1401
Qy      440 strpATrPmetValPheLyGlyGluIleProLyAspGlnTrpMetLyLysTrpTr 460
Db      1402 GTGGAGGTGATGCTGTTTAAAGGGGAAATTCACAAAGACAGTGATGAAAAAGGCTG 1461
Qy      460 pGIuMetLyAspGlnIleValGlyValValGluProValProHIsAspGlnThyTrCy 480
Db      1462 GGAGATGAAGCGAGATAGTTGGGTGGGAAACCTGCCCCCATGATGAACATACG 1521
Qy      480 sAspProAlaSerLeuPheHIsValSerAsnAspTrpSerPheIleArgTyTrThar 500
Db      1522 TGACCCCGCATCTCTGTTCATTTCTATGATTTACTTCATTCATTCGATATTACAGAG 1581
Qy      500 gThrLeuTyGlnPheGlnPheGlnIuAlaLeuCyGlnAlaAlaLysHIsGluGlyPr 520
Db      1582 GACCCCTTACCAATTCAGATTTCAGAGACACTTTGTCAAGCAGCTAAACATGAAGCC 1641
Qy      520 oluHIsLyCyAspIleSerAsnSerThrGlnIuAlaGlyGlnIuLeuPheAsnMetle 540
Db      1642 TCTGCACAAATGGACATCTC -AACTCTACAGAGCTGGGACAAACCTGTCAATTCCT 1700
Qy      540 uArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGlnIuAsnValValGlyAlaLysAs 560
Db      1701 GAGGNTTGAAGAAATCAGAACCCCGAGACCTAGATTGGAAGAAATGTTAGAGCAAGAA 1760
Qy      560 nmetAsnValArgProLeuLeuAsnTyrrPheGluProLeuPheThrTrpLeuLyspG 580
Db      1761 CATGAATGTAAGGCGCTGCTCAACATCTTGAAGCCCTTAATTAACCTGCTGAAGACCA 1820
Qy      580 nAsnIuAsnSerPheValGlyTrpSerThrAspTrpSerProTyrrAlaAspGlnSerI 600
Db      1821 GAACACAAATTTCTTTTGGGATGGAGTACCGACTGAGAGCCATATGCCAACCAAGCAN 1880
Qy      600 elysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrrGluTrpAsnAsp 620
Db      1881 CAAGGTAGAGTAAAGCTTAAATCAGCTTGGAGATGAACATATGAAATGAACAGCA 1940
Qy      620 ngluMetTyrrLeuPheArgSerSerValAlaTyrrAlaMetArgGlnThyPheLeuLysVa 640
Db      1941 TGAATGTACCTGTTCCGATCATTGTTGCAATGATGATGAGGAGGACTTTTAAAGT 2000
Qy      640 llysAsnGlnMetIleLeuPheGlyGluIuAspValArgValAlaAsnLeuLysProAr 660
Db      2001 AAAAATCAGATGATTTCTTTTGGGAGGAGAGTGCAGGTGGCTAATTTGAACCAAG 2060
Qy      660 gIleSerPheAsnPhePheValIthrAlaProLyAsnValSerAspIleIleProArgTh 680
Db      2061 AATCTCCTTAATTTCTTTGCACTGCACCTAAATAAGTGTCTGATTTCAATCTCTAGAAC 2120
Qy      680 rGluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAs 700
Db      2121 TGAAGTTGAAGAGCCATCAGAGATGTCGCCGAGCCGATATCATGATCTTTCCCTCGAA 2180
Qy      700 nAspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPr 720
Db      2181 TGAGACAGAGCTAGAGTTTCTGGGATAGACCAACCACTTGAGACTCTTAACAGCCCC 2240
Qy      720 o 720
Db      2241 T 2241

```

```

ID      AAS21279 standard; cDNA; 3732 BP.
XX
AC      AAS21279;
XX
XX      24-OCT-2001 (first entry)
XX
DE      Human cDNA sequence encoding for PRO1885 polypeptide.
XX
XX      Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW      breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
KW      cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW      adipocyte; A-peptide; factor VIIa; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      M020140466-AZ.
XX
PD      07-JUN-2001.
XX
XX      01-DEC-2000; 2000MO-US32678.
XX
PF      01-DEC-1999; 99MO-US28301.
XX
PR      01-DEC-1999; 99MO-US28634.
XX
PR      02-DEC-1999; 99MO-US28551.
XX
PR      02-DEC-1999; 99MO-US28564.
XX
PR      02-DEC-1999; 99MO-US28565.
XX
PR      09-DEC-1999; 99US-0170262.
XX
PR      16-DEC-1999; 99MO-US30095.
XX
PR      20-DEC-1999; 99MO-US30911.
XX
PR      20-DEC-1999; 99MO-US30999.
XX
PR      30-DEC-1999; 99MO-US31243.
XX
PR      06-JAN-2000; 2000MO-US00227.
XX
PR      06-JAN-2000; 2000MO-US00376.
XX
PR      11-FEB-2000; 2000MO-US00365.
XX
PR      18-FEB-2000; 2000MO-US04341.
XX
PR      18-FEB-2000; 2000MO-US04342.
XX
PR      22-FEB-2000; 2000MO-US04914.
XX
PR      24-FEB-2000; 2000MO-US05004.
XX
PR      24-FEB-2000; 2000MO-US05004.
XX
PR      01-MAR-2000; 2000MO-US05601.
XX
PR      20-MAR-2000; 2000MO-US07377.
XX
PR      21-MAR-2000; 2000MO-US08439.
XX
PR      30-MAR-2000; 2000MO-US08439.
XX
PR      17-MAY-2000; 2000MO-US13705.
XX
PR      22-MAY-2000; 2000MO-US14042.
XX
PR      30-MAY-2000; 2000MO-US14941.
XX
PR      02-JUN-2000; 2000MO-US15264.
XX
PR      10-NOV-2000; 2000MO-US30873.
XX
XX
PA      (GENE ) GENENTECH INC.
XX
XX      Baker KP, Beresini M, Deforge L, Desnoyers L, Flivaroff E, Gao W;
PI      Gerritsen ME, Goddard A, Godowski PJ, Gurney AU, Sherwood S;
PI      Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX      WPI: 2001-408281/43.
XX
XX      P-PsDB: AAU22207.
XX
XX      Isolated, secretory and transmembrane PRO polypeptide used to detect
XX      other PRO polypeptides, link bioactive molecules to cells expressing
XX      PRO polypeptides, and detect the presence of mammalian tumours e.g.
XX      lung, breast, prostate, cervical
XX
XX      Claim 3: Fig 71: 813pp; English.
XX
XX      AAS21244-AAS21518 encode for novel human secretory and transmembrane
XX      PRO polypeptides. The PRO polypeptides are useful to detect other
XX      PRO polypeptides, to link bioactive molecules to cells expressing
XX      PRO polypeptides, to modulate biological activities of cells expressing
XX      PRO polypeptides, and to detect the presence of mammalian lung, colon,
XX      breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX      polypeptide expression in a cell sample to that in a control sample.
XX      Some of the 275 sequences are also useful to stimulate the release of

```

CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 CC
 XX

SO Sequence 3732 BP; 1137 A; 722 C; 821 G; 1052 T; 0 other;

Alignment Scores:

Alignment No.:	Length:	3732
Score:	3720.00	717
Percent Similarity:	83.88%	
Best Local Similarity:	83.76%	
Query Match:	96.15%	
DB:	22	
	Gaps:	1
	Indels:	138

US-09-978-385-2_COPY_19_738 (1-720) x AAS21279 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluasp 20
 DB 94 TCACACCTTATGAGAACAGGCGCAAGACATTTTGGACAAAGTTTAAACACGAAAGCCGAAAGAC 153
 QY 21 LeuPheTyrGlnSerSerIleuAlaSerTyrPheAsnThrAsnIleThrGluGluasp 40
 DB 154 CTGTCTATCAAAAGTTTCACTTCTCTGGAATTTATACACCAATTTTCTGAGAGAAAT 213
 QY 41 ValGlnAsnMetLysAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAACATGAAATATGCTGGGACAAATGCTGCTTTTAAAGAACGTCACAC 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCCAATGATCACTACACAAAGAAATTCACACTGCAAGCTTCAGCTGCAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysTyrLeuAsnThr 100
 DB 334 GCTCTCAGCAAAATGGGCTTCAGGCTCTCAGCAAGCAAGCAAGCGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 394 ATTCTAAATACAAATGAGCAACCACTACAGTACGAAAGTTTGTAAACCAATATATCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaLysSerLeuAspTyr 140
 DB 454 CAAAGATGCTTATATTGACCAAGGTTGAAATGAATGAACAAAGTTTAAAGCTTAC 513
 QY 141 AsnGluArgLeuTyrPheLysTyrGlnSerTyrPheSerGluValGlyLysGlnLeuArgPro 160
 DB 514 AATGAGAGGCTCTGGGCTTGGCAAGCTGAGATCTGAGTGGCGCAAGCAAGCGAGCGCA 573
 QY 161 LeuTyrGlnGluTyrValIleLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluasp 180
 DB 574 TTTATGAAAGAGATGTGTGCTTGAATAATGAGATGCAAGCAAAATATTATGAGGAC 633
 QY 181 TyrGlyAspTyrTyrPheGlyAspTyrGluValAsnGluValAspGlyTyrAspTyrSer 200
 DB 634 TATGGGATTTATTTGAGAGAGAGACTATGAAATGAAGGAGTATGATATGATCAACAC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CCGGGCCAGTATGATGAAATGTGAAACATACCTTTGAGAGATTAACCAATTATATGAA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetLysAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTTCATGCTATGTAGGCGAAAGTTGATGAATGCTATCTCTATATCAAGTCA 813

QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
 DB 814 ATTGATGGCTCCCTGCTCATTTGCTGTGATATGTGGGAGATTTTGGCAAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTTCAGAGTCCCTTGGACGAAACCAACATGATGTTAGTATCATTCATGCTGT 933
 QY 281 AsnGlnAlaTyrPheAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACACAGCTGGGATGACAGAGAAATTCAGAGAGGCGGAGAAATTTCTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPgiLysAsnSerMetLeuThrAspProGlyAsn 320
 DB 994 GGTCTCTTATATATGACTCAAGATTTCTGGAAATTTCTATCTATACGACCCAGGAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAACAGAGTCCCATCCACAGCTTGGAGACCTGGGAGAGCGCACTTCAGAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetLysAspPheLeuThrAlaHisGluMetCysLys 360
 DB 1114 CTATATGTCACAAAGGTGACATGAGAGACTTCTGACAGCTCATCATGATGAGATGGGAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuArgAsnGlyValAsnGlu 380
 DB 1174 ATCCAGTATGATATGATGATGATCTGACAACTTTCTGCTATGAAGAAATGAGCTATATA 1233
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLysSerAlaIleThrProLysHisLeu 400
 DB 1234 GAATTCACAGAGCGTGGGGAATTCATGCTCTTCTGACGCCACCTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIle 420
 DB 1294 AATTCATTTGCTCTTCTGACCCCACTTTTCAAGAGCAATGAACGAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1354 CTGCTCAAAACAGCACTACGATGTTGGAGCTGCGCATTTACTATGATTTAGAGAG 1413
 QY 441 TrrArgTrrMetValPheLysGlyLysIleProLysAspGlnTrrMetLysLysTrrP 460
 DB 1414 TGGAGGTGAGGTCTTTTAAAGGGAAATTCACAAAGCAAGTGGATGATTAAGTGGTGG 1473
 QY 461 GluMetLysArgGlnIleValGlyValAlaGluProValProHisAspGluThrTyrCys 480
 DB 1474 GAGATGAAAGCGAGATATGTTGGGTGGTGGAGCTGCCCCATGATGAACATTAAGTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrrArg 500
 DB 1534 GACCCCGCATCTCTGTCATGTTCTGATGATTCATCATTCATTCATGATTAACAAGG 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1594 ACCCTTTACCAATTTCCATTTCAAGAACCACTTGTCTCAAGCAAGCTTAACGAAGGCCCT 1653
 QY 521 LeuHisLysCysAspLysSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGAGAAACGTTT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCCTCAAAATGTTGAACCTCTCTAGTATTCAGTATTCATTCATTCATTCATTCATTCAT 1772
 QY 537 ----- 537
 DB 1773 TATTTGATTTCTTTGTTCTTAAAGAAATTTTATGGCTCAAAATGCTCATTTTCAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTTAAATTTGTGTGTCAGACAGAACTAGAACATCAACAAATTTGGTGGGCA 1892
 QY 537 ----- 537

Db 1893 CCTTTTTCCTCATATACATACAGACCCCTCTCTCTGTAATTGGAAGGAAGAGCG 1952
 QY 537 ----- 537
 Db 1953 GTTTAGGTGGAATATATCTGTAATATGACATTTCTTCTATCTCCAGAAAGCAATTT 2012
 QY 537 ----- 537
 Db 2013 AGCCAGTCAAGAGGAAGAAACATATGATCATATATATATATGATCATCTGAGACC 2072
 QY 538 ----- 538
 Db 2073 CTTCAAAAGCCCTGAACCCCTTTTGTGTAGCAATATGCTGAGCGCTTGGAATATCA 2132
 QY 546 GlnProThrPheLeuAlaLeuGluAsnValAlaValAlaLysAsnMetAsnValArgPro 565
 Db 2133 GAACCCGTGAGCCCTGAGATGGAAATGTTGTAGAGC -AAGAACATGAATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 Db 2192 CTGCTCAACTACTTGTAGCCCTTATTTACTGCTGGAAGAACAGAAAGATTTCTTT 2251
 QY 586 ValGlyTyrSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 Db 2252 GTGGGATGAGATACCGAGCTGAGTCCATATGACAGACCCAAAGCATCAAGTGAGATAG 2311
 QY 605 rLeuLysSerAlaLeuGluLysAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
 Db 2312 CTTAAATACAGCTCTTGAGATAGATAGCATATGATGAGAACAGACAAATGATGACTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
 Db 2372 CCGATCATCTGTTGCATATGCTATGAGCAGTACTTTTAAAGTAAATAATCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 Db 2432 TCTTTTGGGAGAGATGATGCGAGTGAATGGAACCAAGAAATCTCTTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 Db 2492 CTTTGTCACTGCACCTTAAAAATGCTGTGATATCATCTTCAAGAACTGAATGAAAAAGC 2551
 QY 685 aileArgMetSerArgSerArgIleAsnAspAlaPheArgLysAsnAspAsnSerLeuG 705
 Db 2552 CATCAGATGTCGCGAGACGCTATCATGATGATGCTTCCGTCTGATGACAAACAGCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2612 GTTCTGTGGGATACAGCCACACTTGGACCTCTTAACCCAGCCCT 2657
 RESULT 10
 AA259466
 ID AA259466 standard; DNA; 2262 BP.
 XX AA259466;
 XX 11-APR-2000 (first entry)
 XX Human MPROT15 coding sequence #2.
 DE MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;
 KW heart disease; apoplexy; heart disease; nervous denaturation; ds;
 KW Alzheimer's disease; hormone; cytokine.
 OS Homo sapiens.
 PN JP11318472-A.
 XX 24-NOV-1999.
 XX 22-JAN-1999; 99UP-0014949.

PR 13-MAY-1998; 98GB-0010373.
 PR 18-AUG-1998; 98GB-0018009.
 XX (SMK) SMITHKLINE BEECHAM PLC.
 PA WPI; 2000-109268/10.
 DR MPROT15 polypeptide and MPROT15 polynucleotides - useful for the
 XX treatment of hypertension, myocardial diseases, apoplexy, heart
 XX diseases, nervous denaturation, Alzheimer's disease etc.
 PT Claim 18; Page 15; 22pp; Japanese.
 PS
 XX This is coding sequence #2 of human MPROT15. The MPROT15 polynucleotide
 CC and polypeptide sequences can be used for the treatment of hypertension,
 CC myocardial diseases, apoplexy, heart diseases, nervous denaturation,
 CC Alzheimer's disease and diseases related to the processing of peptide
 CC hormones and cytokines.
 XX
 SQ Sequence 2262 BP; 693 A; 450 C; 523 G; 596 T; 0 other.
 Alignment Scores:
 Pred. No.: 0 Length: 2262
 Score: 3366.50 Matches: 636
 Percent Similarity: 88.58 Conservative: 9
 Best Local Similarity: 88.33 Mismatches: 12
 Query Match: 87.01% Indels: 63
 DB: 21 Gaps: 3
 US-09-978-385-2_COPY_19_738 (1-720) x AA259466 (1-2262)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 88 TCCACCATGAGAGAACAGCCCAAGACATTTTGGCAAGTTTAAACAGAGCCGAAAGC 147
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
 Db 148 CTGTTCTATCAAGTCTCACTTCTCTTCTGGAATTTAAACCATATTTACTGAAGAGAT 207
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 Db 208 GTCCAAAGATGAATATATCTGGGACAAATGCTGCTTTTAAAGAACAGACGACCA 267
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 268 CTTGCCCAATGATATCCATACACAAATTCGAATTCACAGTCAACCTTCAGCTGAG 327
 QY 81 AlaLeuGlnGluAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 328 GCTCTTCACCAAAATGGTCTTCAGTGCCTCTAGAGAACAGCAAGCGTTGAACACA 387
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 Db 388 ATCTTAATATACATATGACACACATACACTACTGAAAAAGTTTGAACCCAGATATCA 447
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 Db 448 CAAAGATGCTTATTTACTTAACACAGGTTGAATGAATGAATGAACAAACAGTTTACACTAC 507
 QY 141 AsnGluArgLeuThrAlaTrpCysLeuSerTrpArgSerGluValAlaGlyLysGlnLeuArgPro 160
 Db 508 AATGAGAGGCTCTGGGCTTGGAAGCTGGAAGCTGAATCTGAAGTGGCAACAGCTGAGGCCA 567
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 Db 568 TTATATGAGAGATGCTGCTTGAATAATGAGATGCAAGAGCAAAATCATATATAGAGAC 627
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 Db 628 TATGGGATATATGGAGAGAGACATATGAATGAATGGGTGATGAGCTATGACTACAGGC 687
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220

PS Claim 10; Page 104-109; 125pp; English.

XX The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-converting enzyme is a zinc metalloproteinase that plays roles in blood pressure regulation and fertility. Zace2 can be expressed by standard recombinant methodology. Zace2 polypeptides are useful for treating an inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis), as targets for identifying modulators of inflammation like arthritis and enterocolitis, as targets for identifying modulators of zinc protease activity, for screening or identifying new angiotensin-converting enzyme (ACE) inhibitors, and as a basis for rational drug design for inhibitory molecules. The nucleic acids can be used to detect the expression of a Zace2 gene in a biological sample, as probes for in vivo diagnosis and for detecting and localizing Zace2 gene expression in tissue samples, to determine whether a subject's chromosomes contain a mutation in the Zace2 gene, and to detect aberrations associated with the Zace2 locus. CC Inhibitors of ACE are used for treating hypertension of various conditions, including left ventricular systolic dysfunction, progressive renal impairment, scleroderma renal crisis, congestive heart failure due to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be used to treat infertility while Zace2 antagonists are used for inducing infertility. The present sequence represents a cDNA encoding the mouse Zace2-5 protein.

XX Sequence 2638 BP; 802 A; 556 C; 611 G; 669 T; 0 other;

Alignment Scores:

Score: 3.9e-313 Length: 2638
Percent Similarity: 3251.00 Matches: 598
Best Local Similarity: 89.06% Conservative: 49
Query Match: 84.03% Mismatches: 73
Gaps: 0
Indels: 0

US-09-978-385-2_COPY_19_738 (1-720) x AAC84368 (1-2638)

QY 1 SerThrIleGluGluGluAlaIysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
DB 160 TCCCTCCACCGAGAAATGCCAAGACATTTTAAACATTTATATCGAAGCTGAAGAC 219
QY 21 LeuPheTyGlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluAsn 40
DB 220 CTGCTTATCAAACTTCACCTTCTTGGATTAATACTAATCACTATGAAAGAAAT 279
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
DB 280 GCCCAAAAGATGAGTGGAGCTGCAGCAAAATGGTCTCTTATGAGAAACAGTCAAG 339
QY 61 LeuAlaGlnMetLysTrpProLeuGlnGluIleGlnAsnLeuThrValLysGlnLeuGln 80
DB 340 ACTGCCCAAGTCTTCTCACTACAGAAATCCAGACTCCGATCATCAAGCCTCAACTACAG 399
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
DB 400 GCCCTTCAGAAAGTGGGTCTTACAGCACTCCAGACAGAAAGAAACAGTTGAGACA 459
QY 101 IleLeuAsnThrMetSerThrIleLysSerThrGlyLysValCysAsnProAspAsnPro 120
DB 460 ATTTGGAACACCAAGACACCATTTACAGTACGGAAGTTTGCAACCCAAAGAACCCA 519
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrp 140
DB 520 CAAAGATGCTTATTCATCTTGAAGCCAGATGGATGAATATGGCAGCAACACAGACTAC 579
QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
DB 580 AACCTGATGCTCTGGGATGGAGGCTGAGGAGCTGAGGTGGCAAGCAGCTGAGCGCG 639
QY 161 LeuTyGlnGluLysValValLeuLysAsnGluMetAlaArgAlaAsnHisLysGluAsp 180
DB 640 TTGATGTAAGAGTATGTGCTACTGAAAGACAGATGGCAAGCAACATTAATACGAC 699
QY 181 TyrGlyAspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyTrpAspTrpSer 200

DB 700 TATGGGATTTATTTGAGAGGGGACTATACAGACAGAGCAAGATGGCTACACTATAC 759
QY 201 ArgGlyGlnIleLeuGluAspValGluIleThrPheGluGluIleLysProLeuTyGlu 220
DB 760 CATAACCGTTATGTAAGATGATACAGCTTCCTCCAGAGATCAACCCATTATAGAG 819
QY 221 HisLeuHisAlaTrpValArgAlaLysLeuMetAlaLysTrpSerTrpIleSerPro 240
DB 820 CATCTCAATGCTATGTGAGGAGAAAGTTGATGATACCTACCTCTCATACAGCCCC 879
QY 241 IleGlyCysLeuProAlaHisLeuLeuLysPheMetTrpGlyArgPheTrpThrAsnLeu 260
DB 880 ACTGATGCTCCCTCCCATTTGGTGTATGTGGGTGAATTTTGGCAAAACCTG 939
QY 261 TyrSerLeuThrValProPheGlyLysProAsnIleAspValThrAspAlaMetVal 280
DB 940 TACCTTTGACGTCTTCCTTCACAGAAACCAACATAGATGTTACTGATGCATATAG 999
QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPheValSerVal 300
DB 1000 AATCAGGGCTGGAGTGCAGAAAGATATTTCAAGACAGCAAGAAATCTTGTCTCTGT 1059
QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
DB 1060 GGCTCTTCATATATACATCAAGATCTGGGCAAACTCTATGCTAGTACAGCAGCAT 1119
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
DB 1120 GGCCGGAAAGTGTGTGCCACCCACAGCTTGATGGAGACAGAGACTTCAGATTC 1179
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
DB 1180 AAGATGTGTCAAAAGTGCATCAATGAGCAACATCTTGACAGCCCAACAGATGGGAC 1239
QY 361 IleGlnTrpAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGluValAsnGlu 380
DB 1240 ATCCATATGACATGACATGCAATGCCAGCACTTCTCTGTAAGAAAGGACCAATGAA 1299
QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
DB 1300 GGCTTCATGAAAGCTGTGGAAGAAATCATGCTATCTTGACAGTCCCAAGCATCTG 1359
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspGlnThrGluIleAsnHe 420
DB 1360 AAATCCATGTTGCTTCTGCCATCCGATTTTCAAGAAAGATGGAAACAGATTAATCTT 1419
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
DB 1420 CTATCTAAGAACGACATGACATTTGTGGAACACTACGCTTACTATGATGTTAGAGAG 1479
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpTrp 460
DB 1480 TGGAGTGTGATGTTCTTGGGGTGAATTTCCAAAGACATGATGAAGAAAGGTGG 1539
QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
DB 1540 GAGATACAGCGGAGATGCTTGTGTGGTGGAGAGCTCTGCTCATGATGAACATACCT 1599
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpArg 500
DB 1600 GACCTGCAATCTGTTTCATGTTCTATGATCTCATCTCATTCATGATATACAGAG 1659
QY 501 ThrLeuTyGlnPheIleGlnGluAlaLeuCysGlnAlaAlaLysHisGluLysPro 520
DB 1660 ACCATTACCAATTCAGATTTCAAGAGGCTTTTGAAGCAGCAGTAAGATATAGTCT 1719
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeu 540
DB 1720 CTGCACAAATGTACATCTCAAAATTTCCATGAGAGCTGGGCAAAAGTGTCTCAAGTCTG 1779
QY 541 ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyValLysAsn 560


```

Db 1780 AGCTTGGAAATTCAGAGCCCTGACCAAGCCTTGAAATGCTGAGAGCAAGAT 1839
QY 561 MetanValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrLeuLysAspGln 580
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1840 ATGGATGTTAAACCACTGCTCAATTAATCCCAACCTGTTGATGCTGCAAGACAG 1899
QY 581 AsnLysAsnSerPheValGlyTrpSerThrAspTyrPheProTyrAlaAspGlnSerIle 600
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1900 AACAGAAATCTTTGTTGGGGTGAACACTGATGACCCCATATGCGACCAAGCAATT 1959
QY 601 LysValArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn 620
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1960 AAGAGAGATTAAGCTTAATAATCAGCTCTTGAGAGTATGATGATGAGCAACAC 2019
QY 621 GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal 640
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2020 GAAATGTTCTCTGCTGCAATCATCTGTCATATGCCATGAGCAAAATATTTTCAATATC 2079
QY 641 LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2080 AAAAACCAAGACAGTTCCTTTCTTACAGAGAGATGACAGATGAGCAATTTGAACCAAGA 2139
QY 661 IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr 680
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2140 GTCTCCTTCTACTTCTTTGTACCTCACCCCAAAATGCTGATGATCTTCTGTAAGT 2199
QY 681 GluValGluLysAlaIleArgMetSerArgSerArgGlnIleAsnAspAlaPheArgLeuAsn 700
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2200 GAAGTGAAGATGATCCATGAGATGCTCGGGGGCCGACATCATGATGCTTGGCCTGCAAT 2259
QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2260 GATTAACAGCTTGAGATTTCTGGGATTCACCCAAACACTTGAGCCACCTTACCAAGCTCTCT 2319
RESULT 12
AAC84370
ID AAC84370 standard; cDNA; 2638 BP.
XX
AC AAC84370;
XX
DT 19-MAR-2001 (first entry)
XX
DE Mouse Zace2-10 protein encoding cDNA.
XX
KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
    zinc metalloproteinase; blood pressure; zinc protease; hypertension;
    ventricular systolic dysfunction; renal impairment; heart failure;
    scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
    antirheumatic; bradykinin inactivator; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 106..2523
FT /tag= a
FT /product= "Zace2-10"
XX
PN WO200070032-A1.
XX
PD 23-NOV-2000.
XX
PF 03-MAY-2000; 2000MO-US11932.
XX
PR 13-MAY-1999; 9905-0311482.
PR 27-AUG-1999; 9905-0384706.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
XX
DR WPI; 2001-025018/03.
XX
P-PSDB; AAB48098.

```

```

PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
PT associated with inflammation such as arthritis and enterocolitis -
XX
PS Disclosure; Page 113-118; 125pp; English.
XX
CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
CC converting enzyme is a zinc metalloproteinase that plays roles in blood
CC pressure regulation and fertility. Zace2 can be expressed by standard
CC recombinant methodology. Zace2 polypeptides are useful for treating an
CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
CC diseases associated with inflammation like arthritis and enterocolitis,
CC as targets for identifying modulators of zinc protease activity, for
CC screening or identifying new angiotensin-converting enzyme (ACE)
CC inhibitors, and as a basis for rational drug design for inhibitory
CC molecules. The nucleic acids can be used to detect the expression of a
CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
CC for detecting and localizing Zace2 gene expression in tissue samples,
CC to determine whether a subject's chromosomes contain a mutation in the
CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
CC Inhibitors of ACE are used for treating hypertension of various
CC conditions, including left ventricular systolic dysfunction, progressive
CC renal impairment, scleroderma renal crisis, congestive heart failure due
CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
CC used to treat infertility while Zace2 antagonists are used for inducing
CC infertility. The present sequence represents a cDNA encoding the mouse
    Zace2-10 protein.
XX
SQ Sequence 2638 BP; 797 A; 557 C; 615 G; 669 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,41e-311 Length: 2638
Score: 3233.00 Matches: 596
Percent Similarity: 89.58% Conservative: 49
Best Local Similarity: 82.78% Mismatch: 75
Query Match: 83.56% Indels: 0
DB: 22 Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x AAC84370 (1-2638)
QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnIleGluAlaGluAsp 20
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 160 TCCCTCACCGAGAGAAATGCCAAGACATTTTAACACATTATACGAGGCTGAAGAC 219
QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnIleThrGluAsn 40
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 220 CTGCTTATATCAAAAGTCACTGCTTCTGGAATTAATATCAATTAAGTGAAGAAAT 279
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerIleAlaPheLeuLysGluGlnSerThr 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 280 GCCCAAAAGATGAGTGAAGGCTGCAGCCCAAAATGCTCTTTATGGAAGAACAGTCTAAC 339
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 340 ACTGCCCAAGTTCTCACTCAACAAATCCAGACTCCGATCATCAAGCTCAACTACAG 399
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 400 GCCCTTCAGCAAAAGTGGTCTTACAGCTCTCAGACAGACAGAAACAAAGTGAACACA 459
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValGlyAsnProAspAsnPro 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 460 ATTCTGAACACCATGACACCATTTACAGTACGTGAAATGCAATTTGCAACCCCAAGACCCA 519
QY 121 GlnGluLysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 520 CAAGATGCTTATTAATTAAGCCAGATTTGATTAATAATGGCGACACACAGACTAC 579
QY 141 AsnGluArgLeuThrPalatPheGluSerTyrPheSerGluValGlyLysGlnLeuArgPro 160
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 580 AACTCTAGAGCTCTGGGATGAGAGGCTGAGAGGCTGAGGTGGCAAGCAGCTGAGAGCGC 639
QY 161 LeuTyrGluGluLysValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180

```

D	b		640	TTGTATGAAGAGTATGTCCTCGTAATAAACGAGATGGCAGAGCAACAATATTAAACGC	699
O	y	181	TyrGlyAspTyrTrpArgGlyAspPyrGluValAsnGlyValAspGlyTyrAspPyrSer	200	
D	b	700	TATGGGGATTATTGGAGAAGGGACATATGAAGCAGAGAGCAGATGGCTACCAACTATAAC	759	
O	y	201	ArgGlyInLeuIleGluAspValLuhStrPhenGluGluIleLysProLeuYrrGlu	220	
D	b	760	CGTACCAGGTGATTTGAAGATGTACAGGTACCTCCGACAGATCAACCATTTGATG	819	
O	y	221	HisLeuHISAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleserPro	240	
D	b	820	CATCTTCATGCCCTATGTAGAGGAGGAAGTGATGGATACCCTTCCTCACATTCAGCCCC	879	
O	y	241	IlegLysLeuProAlaHisLeuLeuGlyAspMetCPGlyArgPhenThrphrAsneu	260	
D	b	880	ACTGGATGCCCTCCCGCCCATTTGCTTGATATGTGGGTAGATTTTGGCAATATG	939	
O	y	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280	
D	b	940	TACCCCTTAGACTGTCCTTCGACAGAAACCAACATAGATGTACTGATCATATGATG	999	
O	y	281	AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal	300	
D	b	1000	AATCGGGCTGGGATGACAAAAGATATTTCAAAGGAGCAGAAAAATTTCTGTCTCTGT	1055	
O	y	301	GlyLeuProAspMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320	
D	b	1060	GGCCTTCCTCATATGACACCAAGATTCGGGGCAACTGTATGCTACACGACACACAT	1115	
O	y	321	ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgile	340	
D	b	1120	GGCCGGAAAGTTGCTGGCACCCCACACTGGGATCTGGGACACGGAGACTTCGAAATC	1175	
O	y	341	LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHISGluMetGlyHis	360	
D	b	1180	AAGATGTGTACAAAGCTACAAATGGACAAACTCTTGACAGCCCATCAGACATGGGACAC	1233	
O	y	361	IlegIntyAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380	
D	b	1240	ATCCAAATGTGACATGGCATATGCCAGGCACCTTTCCTGTAAGAAAGGAGCCAAATGA	1299	
O	y	381	GlyPheHISGluAlaValIGluIleMetSerLeuSerAlaAlaThrProLysHisLeu	400	
D	b	1300	GGGTCCCATGAGACTGTGGAGAAATCATGTCACTTGTGCAAGCTACCCCAAGCATGTG	1355	
O	y	401	LysSerIlegIleLeuLeuSerProAspPheGlnLysAspAsnGluThrGluIleAsnPh	420	
D	b	1360	AAATCCATGTGGCTCTCCSCCATGCCATTTTCAAGAAAGATAGGAAACAGAGATMACTTC	1415	
O	y	421	LeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyMeLeuGluLys	440	
D	b	1420	CTACTGAAACAGCGCTTGTACAAATGGTTGGAAACACTACCGTTACTTACATGTATACAGAG	1479	
O	y	441	TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp	460	
D	b	1480	TGGAGGTGAGTGTCTTTCGGGGGTGAATTTCCCAAGAGCAGTGAATAAAAAGGTGG	1533	
O	y	461	GluMetLysArgGluIleValIGlyValIGluProValProNHASpGluThrTyrcys	480	
D	b	1540	GAGATGAAGCGGAGATGCTGTGGTGTGGAGCCTTCGCTCGATGAAGAAACATACATGT	1599	
O	y	481	AspProAlaSerLeuPheHisValaserAsnAspTyrSerPheIleArgTyrTyThrArg	500	
D	b	1600	GACCTCGATCTCTGTTCCATGATTTCTATATGATTAATCATTCATTCGATATTACCAAGG	1658	
O	y	501	ThreuleuTyrGlnPheGlnPheGlnGluAlaLeuLysGlnAlaAlaLysHISGluGlyPro	520	
D	b	1660	ACCATTTACCAATTCAGATTCACAGAAAGCTTTGTGTCAAGCAGTAAGATATAAGGTCT	1719	
O	y	521	LeuHISLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu	540	

Db	1720	CTGCACCAATGTGACATCTCAAAATTCACACTGAGAGCTGGGACAGAGTTCCTCAAGATGCTG	1779
QY	541	ArgLeuGlyLysSerGluProTrrPThrLeuAlaLeuGluAsnValValGlyAlaLysAsn	560
Db	1780	AGCTCTGAAATTCAGACCCCTGGACCCAGACCTTGAGAAATGTGGTGGGACAGAGAT	1839
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrrLeuLysAspGln	580
Db	1840	ATGAGATGTAACCACTCTCATTTACTTCCAACCGCTTGTTGACTGGCTGAAAGACAG	1899
QY	581	AsnLysAsnSerPheValGlyTrrSerThrAspTrrSerProTyrAlaAspGlnSerIle	600
Db	1900	AACGAAATTTCTTTGTGGGGTGGAAACACTGATAGACCCCATATGCCGACCAAGCAAT	1959
QY	601	LysValAlaGlyIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrrPasnAspAsn	620
Db	1960	AAAGTAGAGATTAAGCCTTAATAATCAGCTTGGAGCTAATTCATTAAGTAATGAGACCAAC	2019
QY	621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640
Db	2020	GAAATGTTCCTGTTCCGATCATCTGTCATATGCCATGAGAAAGTATCTTCAATATC	2079
QY	641	LysAsnGlnMetIleLeuPheGlyGluLysAspValArgValAlaAlaAsnLeuLysProArg	660
Db	2080	AAAAACCGACAGAGTTCCTTCTAGAGGAGAAATGACAGAGAGATTTGAAACCAAGA	2139
QY	661	IleSerPheAsnPhePheValTrrAlaProLysAsnValSerAspIleIleProArgThr	680
Db	2140	GTCGCTCTCTACTTCTTGTGTACCTCACCACCCAAATGTCTGATGTCTATTCCTAGAACT	2199
QY	681	GluValGlyLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700
Db	2200	GAACTGTGAAGATGGCATCAGATGATCTCTGGGGCCGACATCAATGATGTCTTGGGCTGAAAT	2259
QY	701	AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro	720
Db	2260	CATAACAGCCCTGGAGTTCTCTGGGATTCACCCACACATTGAGCCACTTACCAGCTTCT	2319
RESULT 13			
AC	ABK30270/c	ABK30270 standard; cDNA; 2350 BP.	
xx	ABK30270;		
xx	23-APR-2002	(first entry)	
DE	xx	Human G-protein-coupled protease #40.	
DE	xx	Human; ss; gene; G-protein-coupled protease; gene therapy;	
KW	xx	transgenic; protease mediated disorder; proliferative disorder;	
KW	xx	differentiative disorder; developmental disorder;	
KW	xx	haematopoietic disorder;	
OS	xx	Homo sapiens.	
PN	xx	US6331427-B1.	
PD	xx	18-DEC-2001.	
xx	xx		
PE	xx	26-MAR-1999; 99US-0280116.	
xx	xx		
PR	xx	26-MAR-1999; 99US-0280116.	
xx	xx		
PA	xx	(MILL-) MILLENNIUM PHARM INC.	
xx	xx		
PI	xx	Robison KE;	
DR	xx	WPI; 2002-129545/17.	
PT	xx	New polynucleotides encoding protease homologs of the G-protein-coupled	
PT	xx	protease family, useful in identifying agonists and antagonists for	
PT	xx	diagnosis and treatment of protease mediated disorders -	
xx	xx		

Sequence 2350 BP; 603 A; 548 C; 489 G; 710 T; 0 other;

US-09-978-385-2_COPY_19_738 (1-720) x ABK30270 (1-2350)

Dp	1830	AGACTCAGTAAAGAGAGGCTCTGGCGTTGGGAAAGCTGGAGATCTGAGTCCGACAGAGCT	177
QY	158	uAtgPrOleuTyrgIuGlulTyValValLeuLysasnGluMetAlaArgAlaasnIstY	178
Dp	1770	GAGGCCATTATATGAGAGATATGTGGCTTTAAAAAAGAGATGGCAAGCAATCATTA	1711
QY	178	rGuaSPyTyrgIaSPyTyTTPATgGlaSPyTyGlu-ValasnGlyVal-AspGlyTy	197
Dp	1710	TGAGGACTATGGGGATTATTGGAGAGGAGACTTAAAGATTAATGGGGTAAAGTGGCTAT	1651
QY	198	AspTyT-SerArgIyGluLeuLeuLeuLysPvalGluHisThr-PheGluLulLeuLysP	217
Dp	1650	GACTACATGCCCGCGCCAGTTGATTAAGAGATGGAACTATACCTTTTGAAGAGATTAAAC	1591
QY	217	roLeuTyTgIuHisLeuHisAlaTyValAlaArgAlaLysLeuMetAsnAlaTyProSerT	237
Dp	1590	CATTATATATACATCTTATGCTCTATGAGAGGGCAAAATGTAATAGCTATCTCTCT	1531
QY	237	ylLLeSerProLLeGlyCysLeuProAlaHisLeu-LeuGlyAspMetTyrgly-Argph	256
Dp	1530	ATATGAGTCCATTTGGATGGCTCCCTGCTCATTTCCGTTGGTATATGTGGGTTNGATT	1471
QY	256	eTTP-ThAsnLeuTySerLeuThrVal-ProPheGlyGlu-LysProasnIle-Aspy	275
Dp	1470	TTGGGCAAAATCTGACTCTTTGACAGTCCCTTTGACAGCAAAACCAATACAGATG	1411
QY	275	aIThrAspAla-MetValAspGluAlaIThrAspAla-GIArgIlePheLysGluAla-G	294
Dp	1410	TTACGTATGCCATGTGTGGACCAAGCCCTGGAGTACCAAGAAATTTCAAGAGAGCCGG	1351
QY	294	IuLysPhePheValSerValGlyLeuProasnMetThrGluGlyPheTyT-GluasnSer	313
Dp	1350	AGAAATCTTTTATCTGTGTGCTTCTTAATATGACTCAAGAGATCTGGGGAAATTC	1291
QY	314	MetLeuThr-AspProGlyAsnValGluLysAlaValCysHisProThAlaIThrAspLe	333
Dp	1290	ATGCTTACGGGAGCCCGAGAAATGTTCCAGAAAGCAAGTCCCAACCCACAGCTTTGGAGCT	1231
QY	333	u-GlyLysGlyAsp-PheArgIle-LeuMetCysThrLysValIThrMetAspAspPhe	352
Dp	1230	GGGGAGAGGGGACCTTCAGATCTCTTAATGTGCACAAAGTGCACATGTGAGACTTCT	1171
QY	352	uIThrAlaHisSHSGluMetGlyHisIleGluITyrAspMetAlaTyTAlaAlaGluProph	372
Dp	1170	GACACTCATATGATGATGGGGCATCCAGTATGATGGCATATGCTGCACAACTTTC	1111
QY	372	eLeuLeuTyrgsnGlyAlaasnGluGlyPheHisGluAlaValGluITileMetSerLe	392
Dp	1110	TCGTGTAAAGAAATGAGCTTAATGAAGATTCATGAAGAGCGTGGGGAAATATATGCT	1051
QY	392	uSerAlaAlaThrPro-LysHisLeuLysSerIleGlyLeuLeu--SerProaspPheG	411
Dp	1050	TTCTGCACCCACACCTTAAAGCATTTAAATCCCATGTGGTTTTTTTGTCAACCGAATTTTC	991
QY	411	InGluAsnAsnGluIThrGluITileAsnPheLeuLeuLysGlu-AlaLeuThrITrLeuAla	430
Dp	990	AAAGAGACAAATGAACGAAATTAATCTCTGTCTCAAAACAAACACTCCGATTTTGGG	931
QY	431	Thr-LeuProPheIThrTyMetLeuGluLysTyTTPATgTTPMetValPheLysGlyGluIT	450
Dp	930	ACTCTGCCATTACTTACATGTTAGAGAGTGGAGTGGATGCTCTTTAAAGGGGAAT	871
QY	450	eProLysAspGluIThrPheLys-LysITrITrPheLysArgGluITrLeuAlaGlyVal	470
Dp	870	TTCCCAAAACACAGTGGATGAACAAGTGGGGAGATGAAGCGAGAGATGTGGGGCTGG	811
QY	470	al---GluProVal-ProHisAspGluIThrTyCysAspProAlaSerLeuPheHisAla	488
Dp	810	TTGGGAACCTGTGGCCCCCAATGAAGAAATACTGTGACCCCGCATCTCTGTTCATGTT	751
QY	488	SerAsnAspTyTSer-PheITeArgTyTyTThArgIThrLeuTyT-GluPhe-GluPhe	507
Dp	750	TCATATGATATCTCCATTCATTCGATATTTTACAAAGAGACCCCTTTACCCATTCACATTT	691

QY 508 GlnGluAlaLeuGys-GlnAlaAlaLysHISGluGlyProLeuHisLysCysAspIle 527
 DB 690 CAGAGACGACTTTGTCCACACGCTAAACAGAAAGGCCCTTGCACAAATGTGCATCTC 631
 QY 527 r-AsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlu 547
 DB 630 GTAACCTACAGAAAGCTGCAGAAAGCTTCAATATGCTGAGGCTTGAATAATCAGAAC 571
 QY 547 rOTPrThrLeuAlaLeuGlnAsnValValGlyAlaLys-AsnMetAsnValArgProLeu 566
 DB 570 CCGAGACCCCTAGCATGGAAATGTGTAGAGAGCAAGAAAGCAATGTAAGGCCCTG 511
 QY 567 LeuAsn-TyrPheGlu-ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 586
 DB 510 CTCACGACTCTTGGAGCCGCTTATTTACTGCGTGAAGAACCAACAGAAATCTTTTG 451
 QY 586 ALGlyTrpSerThrAspTrpSerProTyrAlaAspGln-SerIleLysValArgIle-Se 605
 DB 450 TGGGATGGAGTACCCAGCTGAGTCCATATGCAGACCAAAACATCAAGTGAAGTAAAG 391
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPhe 625
 DB 390 CCTAAATCAGCTCTTGGAGATTAAGCATATGAAATGACACATGAAATGTCCTGTT 331
 QY 625 eArGserSerValAlaLysAlaMetArgGlnTyrPheLeuLysValLysAsnGln-MetI 645
 DB 330 CCGATCATCTGTGATATGCTATGAGCAGACGACTTTTAAAGTAAATAATCAGCATGA 271
 QY 645 LeuPheGlyGluGln-AspValAlaValAlaAsnLeuLysProArgLysSerPheAsn 664
 DB 270 TTTCTTTTGGGAGAGAGATGTGAGTGGCTTAATTTGAACCAAGATATCTCTTAAT 211
 QY 665 PhePhe-ValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlu 684
 DB 210 TTTCTTTGTGCTGCGCCCTTAAAGTGTCTGATATCTTCTTGAAGTGAAGTTGAAA 151
 QY 684 sAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerIle 704
 DB 150 GCCCATCAGATGTCCCGAGCCGATCAATGATGCTTCCGTCGATGACCAACAGCT 91
 QY 704 uGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 90 AGAGTTTCTGGGATACAGCCAACTTGAGCTCTCTTAACAGCCCT 42
 RESULT 14
 AAC84367
 ID AAC84367 standard; DNA; 2415 BP.
 AC AAC84367;
 AC AAC84367;
 DT 19-MAR-2001 (first entry)
 DE Human Zace2 protein encoding degenerate sequence.
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; human;
 KW antihypertensive; bradykinin inactivator; ds.
 OS Homo sapiens.
 OS Homo sapiens.
 PN WO200070032-A1.
 PN WO200070032-A1.
 PD 23-NOV-2000.
 PD 23-NOV-2000.
 PF 03-MAY-2000; 2000WO-US11932.
 PF 03-MAY-2000; 2000WO-US11932.
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 PR 27-AUG-1999; 99US-0384706.
 PA (ZYMO) ZYMOGENETICS INC.

XX Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 PI WPL; 2001-025018/03.
 DR P-PSDB; AAB48095.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 PS Disclosure; Page 103-104; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a degenerate sequence
 CC encoding the human Zace2 protein.
 CC
 SQ Sequence 2415 BP; 494 A; 218 C; 398 G; 335 T; 970 other;
 Alignment Scores:
 Pred. No: 1.26e-305 Length: 2415
 Score: 3175.00 Matches: 578
 Percent Similarity: 80.39% Conservative: 0
 Best Local Similarity: 80.39% Mismatches: 141
 Query Match: 82.06% Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAC84367 (1-2415)
 QY 2 ThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeu 21
 DB 58 ACNATHGARGARCARCGNNAARACNTTYTNGAYAAATTTAAACAYCARGCGNARGAYTN 117
 QY 22 PheTyrGlnSerSerLeuAlaSerTyrPAsnTyrAsnThrAsnIleThrGluGlnAsnVal 41
 DB 118 TTYTAYCARWSNMSNTYTCNWSNTGAAATAYAAACNAAYATACNARGARAAAYCTN 177
 QY 42 GlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeu 61
 DB 178 CARAAATGAGATAYAGCGNGNGAYAAATGWSGNCNTTYTNAARAKARCARMSNACNTYN 237
 QY 62 AlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAla 81
 DB 238 GGNCAATATGATCCNTYTCNARGARATHCARAAAYTTCNACNTNARNTNCARNTNARGCN 297
 QY 82 LeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIle 101
 DB 298 YTNCAARCARAAAYGGMWSNGNTNYTWNMGANGAYAAARMSNARNGNTNAAAYACNATH 357
 QY 102 LeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAsnProAsnProGln 121
 DB 358 YTNAAIACNATGWSNACNATHTAATWSNACNGNAAAGTNGTAATCCNGATAYCCNCRK 417
 QY 122 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 141
 DB 418 GARTGYTNTNTNTNGARCCNGGNTTAAAGATATATGCGNAAAYMSNTNGATYTAAY 477

QY 142 GIUATGLeuTPAlaITrPGLuSerTrpArgSerGIuValGIuLYSGlnLeuArgProLeu 161
 DB 478 GARMCNTNTGGCGTGGGCGARMSNTGGMGNSNGARNGNGMAARCAATYTMNGCNCNTN 537
 QY 162 TYRGIUGIUTYValValLeuLYSAsnGLuMetAlaArgAlaAsnHisTYrGIuAspTYr 181
 DB 538 TAYGARGARATAGTGTGTNTAARAAAYGARATGGCMNGCNAAYCAATYAGARATAY 597
 QY 182 GIAspTYrTPArgGIAspTYrGIuValAsnGIuValAspGIuTYrAspTYrSerArg 201
 DB 598 GGNAGATYrTGTGGMGNGAGATAYARCTMAAYGNGNGAYGNTATGATTAWSNKN 657
 QY 202 GIuLYSGlnLeuIleGIuAspValGIuHisThrPheGIuLeuIleLYSProLeuTYrGIuHis 221
 DB 658 GGNCArYrTNAHAGARGATNGARCAACNTTYGARARATHAARCCNYTWTATGARCAV 717
 QY 222 LeuHisAlaTYrValArgAlaLYSLeuMetAsnAlaTYrProSerTYrIleSerProIle 241
 DB 718 YTNCAYSCTNAYGTINMGNCNAARNTNAATGAAYCCNATYCCMSNTAYATHTMSNCNATH 777
 QY 242 GIuLYSGlnLeuProAlaHisLeuLeuGIuAspMetTYrPGLuArgPheTYrPThrAsnLeuTYr 261
 DB 778 GGNATGYrTNCNGCNCATYrTYrTNGNGATGTGGGNGMGNNTTYTGACNAATYrTAY 837
 QY 262 SerLeuThrValProPheGIuLYSGlnLYSProAsnIleAspValITrAspAlaMetValAsp 281
 DB 838 WSNATNACNGTNCNTTYGGNCARARCCNAAYrTHAAYGTNACNGAYCGCATGTNGAY 897
 QY 282 GlnAlaITrPAspAlaGlnArgIlePheLYSGlnAlaGlnLYSProPheValSerValGIu 301
 DB 898 CARCMTGGAGAYGCNCARMGNAITTYAARARCCNARARATYrTYrGTNWSNGTNGN 957
 QY 302 LeuProAsnMetITrGlnGIuPheTYrPGLuAsnSerMetLeuThrAspProGIuAsnVal 321
 DB 958 YrTNCNAAYrTAGACNCARGNTTYTGGARAAYSNATGYrTNCNGAYCCNGMAAYGTN 1017
 QY 322 GlnLYSAlaValCYrSHSProThrAlaITrPAspLeuLYSGlnAspPheArgIleLeu 341
 DB 1018 CARARCCNGTNTCYCACCCNACNGTGGATYrTNGNNAARGNGAYrTYrMGNAHYrTN 1077
 QY 342 MetCYSThrLYSValITrMetAspAspPheLeuThrAlaHisLeuMetGIuHisIle 361
 DB 1078 ATGAYACNARAGTNAACNATGAYGAYrTYrTNCNCCNCAAYCARGARATGGNCATATH 1137
 QY 362 GlnTYrAspMetAlaTYrAlaAlaGlnProPheLeuLeuArgAsnGIuLYSAsnGIuLYS 381
 DB 1138 CARATYGAITATGGCNTAYGCGNCARCCNTTYrTNTNGNNAAYGCGCNAAYGARGGN 1197
 QY 382 PheHisGIuAlaValGIuIleMetSerLeuSerAlaIleThrProLYSHSLeuLYS 401
 DB 1198 TTYCAYGARCCNGTNGNGARATHATGWSNYrTMSNCCNCCNCAARCAAYrTNAAR 1257
 QY 402 SerIleGIuLeuLeuSerProAspPheGIuLYSAsnGIuThrGIuIleAsnPheLeu 421
 DB 1258 WSNATHGANTYrTMSNCCNGAYrTYrCARGARGAAYARARCCNARATHAAYrTYrTN 1317
 QY 422 LeuLYSGlnAlaLeuThrIleValGIuTYrLeuProPheTYrMetLeuGIuLYSPro 441
 DB 1318 YrTNAARARCCNTNACNATHGNGNACNYrTCCNTTYrTACNTAYrGTNGARARATGG 1377
 QY 442 ArgITrMetValPheLYSGlnLeuIleProLYSAspGIuThrMetLYSAsnTYrPGLu 461
 DB 1378 MGNrTGTGTNTTYAARGNGARATHCCMAARGAARGATGATGARARARCTGTGGAR 1437
 QY 462 MetLYSArgGIuIleValGIuValGIuProValProHisAspGIuThrTYrCYsAsp 481
 DB 1438 ATGAARMGARATHTGNGNGTNGARCCNGTCCNCAAYCARGARCCNTAYrTGYAY 1497
 QY 482 ProAlaSerLeuPheHisValSerAsnAspTYrSerPheIleArgTYrITrThrArgTYr 501
 DB 1498 CCGCMMSNTNTTYCAYGTNMSNAAYGATAYrTWTATHTGNTATYrTACNMGNA 1557

QY 502 LeuTYrGlnPheGlnPheGlnGlnAlaLeuCYSGlnAlaAlaLYSHSLeuGIuTYrProLeu 521
 DB 1558 YrTNAAYCARrTYrCARrTYrCARGARCCNTYrTNGARCCNCCNARCAAYGARGCNCNTN 1617
 QY 522 HisLYSCysAspIleSerAsnSerThrGIuAlaGIuLYSLeuPheAsnMetLeuArg 541
 DB 1618 CAYARrTYrGAYrTNAHMSNAAYrSACGARGCCNGNCARARATNTTYrTAAYrGTNGMN 1677
 QY 542 LeuGIuLYSLeuSerGIuProTYrPThrLeuAlaLeuGIuAsnValAlaGIuAlaLYSAsnMet 561
 DB 1678 YrTNGNARMSNGARCCNTGACNTYrTNGCAYrTNGARAYrGTNGCNGCNAARAYrANG 1737
 QY 562 AsnValArgProLeuLeuAsnTYrPheGIuProLeuPheThrTYrPheLYSAspGlnAsn 581
 DB 1738 AAYGTNMGNCNTNTNTNAAYrTAYrTNGARCCNTYrTACNCGTYrTNAARGAYCARAY 1797
 QY 582 LYsAsnSerPheValGIuTYrSerThrAspTYrSerProTYrAlaAspGlnSerIleLYS 601
 DB 1798 AARAAYSNTTYrGTNGNGTGSWSNACNGATGSGMSNCNTAYCNCAYCARMSNATHAR 1857
 QY 602 ValArgIleSerLeuLYSLeuAlaLeuGIuAspLYSAlaTYrGIuThrAsnAspAsnGIu 621
 DB 1858 GTMGNATHTMSNTTNAARMSNGNTNGNGAYrAARGATYrGARrTGAAYrGAAYrGAR 1917
 QY 622 MetTYrLeuPheArgSerSerValAlaTYrAlaMetArgGIuTYrPheLeuLYSValLYS 641
 DB 1918 ATGTAYrTNTTYrGNSNMSNGTNGCNTAYrGCATGTMGNCARATYrTYrTNAARGTNAAR 1977
 QY 642 AsnGlnMetIleLeuPheGIuLYSGlnAspValArgValAlaAsnLeuLYSProArgIle 661
 DB 1978 AAYCARATGATHTNTTYrGNGARGARGAYrTNGNGTNGCNAAYrTNAARCCNMGNAH 2037
 QY 662 SerPheAsnPhePheValITrAlaProLYSAsnValSerAspIleLeuProArgTYrGIu 681
 DB 2038 WSNrTYrTAATYrTYrTNGNCCNCCNAAARAGTMSNGATrTHAATHTCCNMGNCNGAR 2097
 QY 682 ValGIuLYSAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAsp 701
 DB 2098 GTNGARARCCNATHMGNTGMSNMGNSMGNAATrHAAYrGCNTTYrTNGNTNAAYrGAY 2157
 QY 702 AsnSerLeuGIuPheLeuGIuIleGlnProThrLeuGIuTYrProAsnGlnProPro 720
 DB 2158 AAYWSNTYrGATrTYrTNGNATHCARCCNACNTYrTNGCNCNAAYCARCCNCCN 2214
 RESULT 15
 AAS42515
 ID AAS42515 standard; cDNA; 3474 BP.
 AC AAS42515;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human cDNA encoding an mdt protein, clone LI:347572.1:2000MAY01.
 KW Human; molecules for disease detection and treatment; mdt; ss;
 KW Antiartherosclerotic; hepatotropic; antiproliferative; cytosolic;
 KW Immunosuppressive; antidiabetic; antiproliferative; neuroprotective;
 KW osteopathic; antidiabetic; antiproliferative; cell proliferative disorder;
 KW arteriosclerosis; cirrhosis; poriasis; cancer; adenocarcinoma;
 KW leukemia; breast cancer; autoimmune disorder; AIDS;
 KW acquired immunodeficiency syndrome; Addison's disease;
 KW diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.
 XX
 OS Homo sapiens.
 XX
 PN WO200162922-A2.
 XX
 PD 30-APR-2001.
 XX
 PE 21-FEB-2001; 2001WO-US05896.
 XX
 PR 24-FEB-2000; 2000US-0185213.
 PR 16-MAY-2000; 2000US-0205232.

[illegible]

Db 1530 TGAAGAGTGTGGGAGATGACCAAGCAATAGTTGGGTGTGTGGAACTGTGCCCC 1589
 QY 475 IsaspGlnThrTyr-CysaspProAlaSerLeupheHisValSeranspTyrSerphe 494
 Db 1590 ATGATGAACATATCTGTGACCCCGCATCTCTGTCCATGTCTATATGATCTATCT 1649
 QY 495 IleArgTyrThrThrArgThrLeu-TyrGlnPheGlnPheGln-GluAlaLeu-CysGln 513
 Db 1650 ATTGGATATTACACCAAGACCCCTGTACCAATTCAGATTCCAAAGACCTTTGTCAA 1709
 QY 514 AlAlaLysHisGlnGluGlyProLeuHisLys-CysAspIle-SerAsnSerThrGlu---A 532
 Db 1710 GCAGCTAAACATGAGAGCCCTCTGCACAAATGTGTGACATTTCTAAATTTCTACGAACGTC 1769
 QY 532 IacGlnLys-LeupheasMetLeuArgLeuGlyLys-SerGluProThrLeuAl 551
 Db 1770 GTGGACAGACACTGTTCATATGTGAGGCTTGAAACTGCAAACTGAGACCTGAC 1829
 QY 551 AlauGluAsnValVal-GlyAlaLysAsnMetAsnValArgPro-LeuLeuAsnTyrPhe 570
 Db 1830 ATTGGAAATGTTGTAGGACCAAGCAATGATGTAAAGCCACTGTCAACTACTT 1889
 QY 571 GluProLeupheThrThrLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThr 590
 Db 1890 GAGCCCTTATTTACCTGTGCTGAAGACCAAGCAAAATTTCTTTGTGGATGAGTACC 1949
 QY 591 AspTrpSerProTyrAlaAspGlnSerIle-LysValArgIleSerLeuLysSerAla 610
 Db 1950 GACTGGAGTCCATATGACGACCAACAGCATCACAAGTGAAGATTAAGCTTAATACAGCTCT 2009
 QY 610 uGly-AspLysAlaTyrGluThrPheAsnAsp-AsnGluMetLysLeupheArgSerSer-Val 629
 Db 2010 TGGCAGATTAAGATGATGAAATGAGACCAAGCAAAATGATGCTGTCCATATCTGTG 2069
 QY 629 LAlaTyrAla---MetArgLysThrPheLeu-LysValLysAsnGlnMetIleLeupheG 648
 Db 2070 TGGATATTGTTAATGTGAGGAGTACTTTTAAACAGTAAATAATTCAGATATCTTTTG 2129
 QY 648 LysGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPheValT 668
 Db 2130 GGGGAGGAGATGTGCGAGTGGCTTAATTTGAAACCAAGATCTCTTAATTTCTTTGTC 2189
 QY 668 hrAlaProLysAsnValSer-AspIleIleProArg-ThrGluValGluLysAlaIle 687
 Db 2190 CTGACCTAAATATGTCTGTGATATCTTCTAGAAATGGAAGTGAAGGACCATCAG 2249
 QY 687 gmetSerArgSerArg-IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPhe 707
 Db 2250 GATGTCCCGGAGCGGCTGATCTCATGATGCTTCCGTGCAATGACGACGCTAGAGTTTC 2309
 QY 707 euGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2310 TGGGATATACACCAACACTTGGACCTCTTAACCAAGCCCTT 2350
 RESULT 16
 AAC84369
 ID AAC84369 standard; DNA: 2415 BP.
 AC AAC84369;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Mouse Zace2-5 protein encoding degenerate sequence.
 XX
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;
 KW ventricular systolic dysfunction; renal impairment; heart failure;
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;
 KW antiarthritic; Bradykinin inactivator; ds.
 XX
 OS Mus sp.
 XX
 PN W0200070032-A1.

XX
 PD 23-NOV-2000.
 XX
 PF 03-MAY-2000; 2000MO-US11932.
 XX
 PR 13-MAY-1999; 99US-0311482.
 PR 27-AUG-1999; 99US-0384706.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;
 XX
 DR WPI, 2001-025018/03.
 DR P-PSDB; AAB48097.
 XX
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases
 PT associated with inflammation such as arthritis and enterocolitis -
 XX
 PS Disclosure; Page 112-113; 125pp; English.
 XX
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood
 CC pressure regulation and fertility. Zace2 can be expressed by standard
 CC recombinant methodology. Zace2 polypeptides are useful for treating an
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),
 CC diseases associated with inflammation like arthritis and enterocolitis,
 CC as targets for identifying modulators of zinc protease activity, for
 CC screening or identifying new angiotensin-converting enzyme (ACE)
 CC inhibitors, and as a basis for rational drug design for inhibitory
 CC molecules. The nucleic acids can be used to detect the expression of a
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and
 CC for detecting and localizing Zace2 gene expression in tissue samples,
 CC to determine whether a subject's chromosomes contain a mutation in the
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.
 CC Inhibitors of ACE are used for treating hypertension of various
 CC conditions, including left ventricular systolic dysfunction, progressive
 CC renal impairment, scleroderma renal crisis, congestive heart failure due
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be
 CC used to treat infertility while Zace2 antagonists are used for inducing
 CC infertility. The present sequence represents a degenerate sequence
 CC encoding the mouse Zace2-5 protein.
 XX
 SQ Sequence 2415 BP; 488 A; 228 C; 397 G; 324 T; 978 other;
 Alignment Scores:
 Pred. No.: 2,19e-252 Length: 2415
 Score: 2640.00 Matches: 477
 Percent Similarity: 71.838 Conservative: 38
 Best Local Similarity: 66.538 Mismatches: 202
 Query Match: 68.234 Indels: 0
 DB: 22 Gaps: 0
 US-09-978-385-2_COPY_19_738 (1-720) x AAC84369 (1-2415)
 QY 4 GluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeupheTyr 23
 ||||| ||||| : : ||||| ||||| ||||| |||||
 Db 64 GARGAARAYGCNNAARACNTTYYTNAAYTAAYTAAAYCARGARGCARGAYTTNMSAR 123
 QY 24 GlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGlnAsn 43
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 124 CARMSNWSNYTNGCWNSTGGAAYTAAYTAACNAAYTAHACGAGAGAAAGCCNCARAR 183
 QY 44 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuGluGlnSerThrLeuAlaGln 63
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 184 ARGWSMGARGCNGCNCNNAARTRGMSNGCNTTYYTARGARGACARMSNAARCCNCAR 243
 QY 64 MetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeuGln 83
 :
 Db 244 WSNWTYWSNYTNCARARATHCARACNCCNATHAARMGNCARAYTNCARCNNTNCAR 303
 QY 84 GlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsn 103
 ||||| ||||| : : : : : : : : : : : : : : : :

Dd	304	CARMSNGNWSMWSGCVYTWMSNCCNGAVAAARAAYAAARCAATYTAAYACNAHTVTNMAV	363
Qy	104	ThrmSetSerrITLeTySerThrclyLysValCysAsnProAspAsnProGlnGlyCys	123
Dd	364	ACNATGWSNACNATWTATWMSNACNGNBARGNTGTAAAYCCNMAAFAAACCCNBARARFTG	423
Qy	124	LeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGluArg	143
Dd	424	YTNNTNTNGARCCCGGNTYTNQAYARATHTATGGCNACNWSNACNGATTAATAAYASNMG	483
Qy	144	LeuTrpAlaTyrGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlu	163
Dd	484	YTNNGCGGNTGGARGCNTGCGMGNCNCNARGNNGNNAARCAATYTMGNCNTNTNAYGAR	543
Qy	164	GluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGlyAsp	183
Dd	544	GARTAYGNTGNTYTNAARAAGAARATGGCNCMGNCNMAAYATTAAYAAAGATYAYGNGAY	603
Qy	184	TyrTrpArgGlyAspTyrGluValAsnGlyValAsnGlyTyrAspTyrSerArgGlyGln	203
Dd	604	TATYTGMGNGGGAATTAAGARGCNGARGGNCNGAYGNTYAAATAAAYAAWGNAAAYAR	663
Qy	204	LeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeuHis	223
Dd	664	YTNATHGARGAYGTNGARMGNACNTTTCGNCARGARHAARCCNTYTATAGARCAATYNQAY	723
Qy	224	AlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGlyCys	243
Dd	724	GCNTAYGTGNNMGNAARTYTNAATGAYACNTAACCMWNTAVATHTMSNCCNACNGNTGY	783
Qy	244	LeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSerLeu	263
Dd	784	YTNCCNGSCNAYTNTYNGNGAATATGTGGGGNGMNTYTGACNAAATYTNTAYCCNTN	843
Qy	264	ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAsnGlnAla	283
Dd	844	ACNGNCCNTTGGCACAARACCNAAAYTHGAYGTNACNGAYGCATGATGATGAAYCARGN	903
Qy	284	TrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuPro	303
Dd	904	TGGGAYGCGNARGNNAHTTTCARARGCNGARARATYTTYGTWMSGNTGNTNCCN	963
Qy	304	AsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLys	323
Dd	964	CATATGACNARGGNTYTTGGCNAAYSNATGYTNACNGARCCGCGAYGGNMGNAR	1023
Qy	324	AlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCys	343
Dd	1024	GTGNTGNTGAYCCACNACGNTGGGAYTTNGNCAVGGNGAYTTMGNATHAARATGTGT	1083
Qy	344	ThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyr	363
Dd	1084	ACNNAARGTNACNATGAAATYTTYTTCNCGNCNCAVYGARATGGGCACATHTCARATY	1143
Qy	364	AspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHis	383
Dd	1144	GAYATGCGNTYAGCMGNCARCCNTYTTNTNMGAAYGGGCAAAAGARGGNTYCAV	1203
Qy	384	GluAlaValGlyGluIleMetSerLeuSerIleAlaIleTrpProLysHisLeuLysSerIle	403
Dd	1204	GARGCNGTNGNGARATATATGHSNTYTNMSNCGNCNACNCCNNAARCAATYTNAARHSNATH	1263
Qy	404	GlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLys	423
Dd	1264	GGNTYTNCCNMSNGAYTTYCARGARGAYMSNGARACNGARATHAATYTYTNTNNAAR	1323
Qy	424	GlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArgTrp	443
Dd	1324	CARGCNTYNAACNATHTGNGNAACNTTCCNTTYACNTYATGTYTGAAARATGGCMGNTGG	1383
Qy	444	MetValPheLysGlyGluIleProLysAsnGlnTrpMetLysLysTrpTrpGluMetLys	463
Dd	1384	ATGGNTYTMGNGGARATHTCCNNAARCARCAFTGGATGAARMAATGGTGGARRTTGAR	1443

QY	464	ArgGluLeuValaIGlyValaGluProValaProHisAspGluGluThrTyrCysAspProAla	483		
Db	1444	MGNGRARATHGNGNGNGTGNGTGNGARCCNTTCCNCATGAYGRACNATATGTGATCCGCNC	1503		
QY	484	SerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrArgTThrLeuTyr	503		
Db	1504	MSNTNTTTCAYAGTWSMAAYGAATAYMSNTTATVTHMGNTATVTAACMGMAACNATHPAY	1563		
QY	504	GlnPheGlnPheGlnGlnGluAlaLeuCysGlnAlaAlaAlaLysHisGluGlyProLeuHisLys	523		
Db	1564	CARTTYCARTTYCAGARGCAGCYNTTGTGCARCCNCNARRTATYAAAGGWSNYTTCACAAAR	1623		
QY	524	CysAspIleSerAsnSerThrGluAlaGlnLysLeuPheAsnMetLeuArgLeuGly	543		
Db	1624	TGYGATATHSMAAYWSNACNGARGCNGCNCARARRTNTNTNARATGTYTWSYTTGNG	1683		
QY	544	LysSerGluProTyrPThrLeuAlaLeuGluAsnValaGlyAlaLysAsnMetAsnVal	563		
Db	1684	AAWMSNGARCCNTGGACNAARGCUNTNGARAAAYGNGTNGCNGMGMAAATGAGVGTN	1743		
QY	564	ArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsn	583		
Db	1744	AARCCNTNTYTNAAATYATTYTCARCCNTNTTTCAYTGGYTNARGARCAAAATGMAAY	1803		
QY	584	SerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIleLysValArg	603		
Db	1804	WSNTTGTGNGTGGAAATACNGARTRGWSNCCNTATGCGNGATYCARWSNATHAAGTMMGN	1863		
QY	604	IleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPAsnAspAsnGluMetTyr	623		
Db	1864	ATHMSNTYTNARWSNCGCYNTNGCGNCNAAYGCNTATYGARTGACNAAYAAAGARATGTY	1923		
QY	624	LeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGln	643		
Db	1924	YTNTTTCMGMSNWSNGTNGCNTATGACNTTGGNARFATYTTTWSNATHAATAAAACAR	1983		
QY	644	MetIleLeuPheGlyGluGluAspValaArgValaAlaAsnLeuLysProArgIleSerPhe	663		
Db	1984	ACNGTNCNTTYYTNGARGARGAGTNGTNGTNTMSNGATYTNARCCMGNGTWSNTTY	2043		
QY	664	AsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTThrGluValaGlu	683		
Db	2044	TAYTYYTYYTNGACWSNCCNCAAAAYGTNMSNAGYTNATHCCMGMSNGARGTNGAR	2103		
QY	684	LysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer	703		
Db	2104	GAYGCNATHHMGNAATGWSNMGNGMNGNTTHAAYGATGNTTYGGNTNAAAGAAATWSN	2163		
QY	704	LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720			
Db	2164	YTNGARTYYTNGGNATHCAACNACNATNARCCNCTATYCARCCNCCN 2214			
RESULT 17					
AC	AAQ10328	standard; DNA; 2477 BP.			
XX	AAQ10328;				
AC	AAQ10328;				
XX	10-APR-1991	(first entry)			
DE	Encodes human testicular angiotensin conversion enzyme.				
XX					
KW	human testicular angiotensin conversion enzyme; tACE;				
XX	male sterility; ss.				
OS	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FT	CDS	29..2227			
FT	/tag=a				
FT	/product= human tACE				
FT	Peptide 29..91				

QY 492 TyrSerPheIleArgTyrThrArgThrIleuTyrGlnPheGlnIleuAlaLeu 511
 Db 1673 GTGCGTATACATGAGTATGTCAGCTTCATCATGTCAGTCCACGAGCGACG 1732
 QY 512 CysGlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 531
 Db 1733 TGGCAGGACGCTGGCCACACGAGCGCCCTGCACAACTGTGACATCTACACAGTCAAGAG 1792
 QY 532 AlAGlYlInLysIlePheAsnMetIleArgLysGlySerGlnProThrPheAla 551
 Db 1793 GCGCGGACGCGCTGGCCAGACCGCCATGAGCTGGCTTCAGTGGCCGCGCGGAGCC 1852
 QY 552 LeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLysAsnTyrPheGlu 571
 Db 1853 ATGACGCTGATCACAGGCGCCACCAACATGACGCGCTGGCCATGTGTGAGCTACTTCAAG 1912
 QY 572 ProLeuPheThrIleuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 588
 Db 1913 CCGCTGCTGAGTGGCTCCGACGAGAGACGAGCTGCATGGGAGAAAGCTGGGCGCG 1972
 QY 589 SerThrAspTyrSerProTyrAlaAspGlnSer 599
 Db 1973 CAGTACACATGAGACGCGGACCTCGCGCTCA 2005
 RESULT 18
 AAA38330 standard; DNA; 4020 BP.
 AC AAA38330:
 XX
 DT 21-AUG-2000 (first entry)
 XX
 DE Human angiotensin-converting enzyme (ACE) coding region.
 XX
 KW Angiotensin-converting enzyme gene; ACE; coding region; polymorphism;
 KW polymorphic marker; cardiovascular disease; myocardial infarction;
 KW unstable angina; hypertension; atherosclerosis; stroke; prognosis;
 KW drug screening; treatment outcome; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN MO200022166-A2.
 XX
 PD 20-APR-2000.
 XX
 PE 13-OCT-1999; 99WO-IB01678.
 XX
 PR 14-OCT-1998; 98US-0104286.
 PR 14-OCT-1998; 98US-0104302.
 PA (EURO-) EURONA MEDICAL AB.
 PI Norberg LT, Andersson MK, Lindstrom PHR, Jonsson L;
 WP1; 2000-318010/27.
 PT Assessing cardiovascular status in humans involves comparing test
 PT polymorphic pattern comprising polymorphic positions within genes
 PT encoding specific proteins, with reference polymorphic pattern -
 PS Disclosure; Page 114-115; 126pp; English.

CC cardiovascular disorders such as myocardial infarction, unstable angina,
 CC hypertension, atherosclerosis and stroke. They are also useful for
 CC predicting the likely cardiovascular status of a patient given a
 CC treatment regimen comprising administration of cardiovascular drugs
 CC (e.g., ACE inhibitors, beta-adrenergic receptor antagonists (beta-
 CC blockers) or calcium channel blockers). One or more polymorphic markers
 CC provides a basis for predicting the outcome of a treatment regimen.
 CC Fragments of the genes comprising a polymorphic site may be used as
 CC primers and probes for detecting genetic polymorphisms or in molecular
 CC library arrays for high throughput screening. The genes, and the proteins
 CC they encode are useful in the screening of potential cardiovascular
 CC drugs. Determination of an individual's polymorphic pattern reduces or
 CC eliminates trial and error in selecting a treatment for a particular
 CC individual cardiovascular patient. It also provides the ability to
 CC eliminate patients from clinical trials who are predicted to be
 CC non-responsive, or at a risk for an adverse response, to a particular
 CC treatment regimen. Adverse results in an early trial can be evaluated to
 CC identify polymorphic patterns so that the adverse results can be
 CC correlated with a sub-population of the test population, permitting
 CC exclusion of such sub-populations from the treatment group. Beneficial
 CC drugs can be approved for use in the appropriate population, thereby
 CC decreasing the number of patients required for a clinical trial, which in
 CC turn decreases the duration and cost of such trials. Sequences A38328 and
 CC A38330 represent, respectively, intron 16 and the coding region of
 CC the human ACE gene (Genbank X62855, J04144). The polymorphic sites
 CC identified are 375A/C, 582C/T, 731A/G, 1060G/A, 1215C/T, 2136G/A,
 CC 2328A/G, 2741G/T, 3387T/C, 3503G/C, 3906G/A; and a deletion of
 CC nucleotides 1451-1783 in intron 16.
 XX
 SQ Sequence 4020 BP; 857 A; 1261 C; 1174 G; 728 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,13e-122 Length: 4020
 Score: 1337.00 Matches: 255
 Percent Similarity: 61.05% Conservative: 118
 Best Local Similarity: 41.73% Mismatches: 204
 Query Match: 34.56% Indels: 34
 DB: 21 Gaps: 9
 US-09-978-385-2_COPY_19_738 (1-720) x AAA38330 (1-4020)
 QY 2 ThrIleGluGlnIleAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsnLeu 21
 Db 1952 ACTGATGAGGCTGAGGCCACGAACTTGTGAGAGAAATGACCGGACNATCCAGGTGTG 2011
 QY 22 PheTyrGlnSerSerIleuAlaSerTyrPheAsnTyrAsnIleThrGluGlu----- 39
 Db 2012 TGAACGAGTATGCGAGGCGCAACTGGAATCAACACCAATCAGACAGAGACGAC 2071
 QY 40 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerIle 53
 Db 2072 AAGATTGCTGTCAGAGACATGCAATATGACCAACACACC----- 2113
 QY 54 PheLeuLysGluGlnSerThrIleuAlaGlnMetTyrProLeuGlnIleGlnAsnLeu 73
 Db 2114 -----CTGAAGTATAGGACACCGACGACGAGAAAGTTGATGTGACCAATTCAGACAC 2167
 QY 74 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 93
 Db 2168 ACTATCAACGCGATCATAAAGAGGTTCAGGACCTAGAAAGCGGCGAGCCTGCTGCCG 2227
 QY 94 LysSerLysArgLysAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 113
 Db 2228 GAGCTGAGAGAGTACAAACAGATCTGTGTGATATGAAACCACTACAGGCTGCCACT 2287
 QY 114 ValCysAsnProAspAsnProGlnIleLysLeuLeuGluProGlyLeuAsnGluIle 133
 Db 2288 GTGTGACACCGAATGCG-----AGCTGCGTGCAGCTGAGCCAGATCTACAGAAATG 2341
 QY 134 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalatyrPgluSerTyrArgSerGlu 153
 Db 2342 ATGGCCACATCCCGGAAATATATAGAACCTGTATGGCATGTGAGGCGTGGCGAGACAG 2401

154 ValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMetAla 173
 2402 GGGGGAGAGCCATCTCTGAGTTTACCCGAATATGCTGGAAGTCAATCAACAGAGCTGCC 2461
 174 ArgAlaAsnHisTyrGlnLysPheTyrGlnLysPheTyrGlnLysPheTyrGlnLysPhe 193
 2462 CGGCTCAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2521
 194 ValAspGlyTyrAspTyrSerArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 213
 2522 GTGGAG-----CAAGACCTGGAGGGCTCTCCAG 2551
 214 GluLeuLysProLeuTyrGluHisLeuHisAlaTyrValAlaGluLysLeuMetAsnAla 233
 2552 GAGCTGAGAGCCACTTACCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 2611
 234 Tyr---ProSerTyrLysPheProLysLeuPheProAlaHisLeuLeuGlyAspMet 252
 2612 TACGGGGCCAGCATCACTGAGGGGCCATCTCTGCTCACTCTCTGCTGCTGCTGCTGCTGCTG 2671
 253 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 272
 2672 TGGGGGAGAGCTGCTGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2731
 273 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgLysPheGlu 292
 2732 ATGGACACACAGAGGCTATCTAAAGAGGCTGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAG 2791
 293 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 312
 2792 GCTATGATTTCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3261
 313 SerMetLeuThrAspProGlyLysValGlnLysAlaValGlnLysPheProThrAlaTrpAsp 332
 2852 TCGATGCTGGAG 2911
 333 LeuGlyLysGly---AspPheArgLysLeuMetCysThrLysValThrMetAspAspPhe 351
 2912 TTCTACACAGGAG 351
 352 LeuThrAlaHisGlnMetGlnHisLysLysLysLysLysLysLysLysLysLysLysLysLys 371
 2972 GTGGTGGCCCAACCAATGAGGAG 3031
 372 PheLeuLeuArgAsnGlnAlaAsnGlnLysPheHisGlnAlaValGlnLysLysLysLysLys 391
 3032 GTGGGCTTGGAG 3091
 392 LeuSerAlaAlaThrProLysHisLysLysLysLysLysLysLysLysLysLysLysLysLys 411
 3092 CTCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACT 3151
 412 GluAspAsnGluThrGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 431
 3152 AGGAGC---GAGCATGACATCAATCTTCTGATGATGATGATGATGATGATGATGATGATGAT 3208
 432 LeuProPheThrTyrMetLeuGlnLysTyrArgTyrMetValPheLysGlyGluLeuPro 451
 3209 ATCCCTTCACTGCTCACTGCTCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3268
 452 LysAspGlnTrpMetLysLysTyrTrpGlnMetLysArgGlnLysValGlnValGln 471
 3269 AAGGAGAACTATTAACAG 3328
 472 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheLysValSerAsnAsp 491
 3329 CCAAGTGGCCAGACATCAAG 3388
 492 TyrSerPheLysGlyTyrThrArgThrLeuTyrGlnPheGlnPheGlnLysLysLysLysLys 511
 3389 GTGGCTTACATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3448
 512 CysGlnAlaAlaLysHisGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 531

3449 TCCAGGAGAGCTGGCCACAGCGGCCCTGACAAAGTGTGACATCTACAGTCCAGTCCAGAGAG 3508
 532 AlaGlyLysLysLeuPheAsnMetLeuArgLeuGlnLysSerGluPheProThrLeuAla 551
 3509 GCCGGGAGAGCCCTGGGAG 3568
 552 LeuLysAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
 3569 ATGCACTGATGACAGGAG 3628
 572 ProLeuPheThrTyrPheLysAspGlnAsnLys-----AsnSerPheValGlyTyr--- 588
 3629 CCGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3688
 589 SerThrAspTyrSerProTyrAlaAspGlnSer 599
 3689 CAGTACAACTGAG 3721
 Db
 RESULT 19
 AAK9395
 ID AAK9395 standard; DNA: 4022 BP.
 AC
 XX AAK9395:
 XX
 XX 27-JUN-2002 (first entry)
 DE
 XX DNA of APP related human homologue hcp51674.
 KW
 KW Neuroprotective; neurotrophic; transgenic fly; Alzheimer's disease; Abeta;
 KW amyloid precursor protein; tissue-specific expression control; human APP;
 KW APP pathway modulator; gene therapy; gene; ds.
 OS
 XX Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT CDS 148..3945
 FT /tag- a
 FT /product- "Protein of human homologue hcp51674"
 FT /note- "No start codon"
 XX
 PD WO200226820-A2.
 XX
 XX 04-APR-2002.
 XX
 PF 01-OCT-2001; 2001MO-BP11345.
 XX
 PR 29-SEP-2000; 2000US-236893P.
 PR 14-JUN-2001; 2001US-298309P.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M,
 PI Reinhardt MMH, Zisman S;
 XX
 DR WPI: 2002-315796/35.
 DR P-PSDB: AAO20501.
 XX
 PT New transgenic fly, containing DNA encoding an Abeta portion of human
 PT APP, useful for identifying agents which modulate the APP pathway and
 PT which can be used to treat Alzheimer's disease -
 PS
 PS Example 4; Page 93-94; 129pp; English.
 CC
 CC The invention relates to a transgenic fly whose genome comprises DNA
 CC encoding a polypeptide having the Abeta portion of human amyloid
 CC precursor protein (APP), fused to a signal sequence. The DNA sequence
 CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
 CC the specification. The DNA sequence is operably linked to a tissue-
 CC specific expression control sequence. Expression of the sequence gives
 CC the fly an altered phenotype. The purpose of the invention is for
 CC identifying agents that inhibit or promote the expression and/or function

XX AA004027;
 XX 07-SEP-1990 (first entry)
 XX Human angiotensin converting enzyme (ACE) gene.
 XX human angiotensin converting enzyme; hypertension;
 XX bradykinin;
 XX synthetic.
 XX Key Location/Qualifiers
 XX sig_peptide 23..51
 XX CDS /*cag- a
 XX /*tag- 23..3944
 XX /*tag- b
 XX MO9003435-A.
 XX 05-APR-1990.
 XX 29-SEP-1989; 89WO-FR00469.
 XX 27-SEP-1988; 88FR-0012620.
 XX (INRM) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
 XX Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P;
 XX WPI: 1990-058128/08.
 XX P-PSDB; AAR04111.
 XX New DNA encoding human angiotensin converting enzyme used eg in
 XX diagnosis of hypertension, evaluation of enzyme inhibitors
 XX Disclosure; ; p: French.
 XX This sequence encodes pre-ACE. In the mature protein the 29 amino acid
 XX signal peptide is absent. ACE hydrolyses angiotensin I and/or kinins,
 XX most notably bradykinin. Vectors contg. the given sequence are used to
 XX produce ACE for therapeutic use eg to hydrolyse kinins implicated in
 XX inflammation.
 XX Sequence 4024 BP; 857 A; 1263 C; 1175 G; 729 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,13e-122 Length: 4024
 Score: 1337.00 Matches: 255
 Percent Similarity: 61.05% Conservative: 118
 Best local Similarity: 41.73% Mismatches: 204
 Query Match: 34,56% Indels: 34
 DB: 11 Gaps: 9
 US-09-978-385-2_COPY_19_738 (1-720) x AA004027 (1-4024)
 QY 2 ThrIIGLGLUGLInAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 21
 DB 1952 ACAGTATGAGCTAGGCGACGCAAGTTGTGAGGAATATGACGACATCCCGAGTGTG 2011
 QY 22 PheTyrGlnSerSerLeuAlaSerTPasTyrAsnThrAsnIleThrGluGlu----- 39
 DB 2012 TGGAGAGATATGCCGAGCGCACTGAGACTACACACCACTCCACACAGAGACCAGC 2071
 QY 40 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAla 53
 DB 2072 AAGATTCTGCTGACAGAGAACATGCAAAATAGCCAACACACACC----- 2113
 QY 54 PheLeuLysGlnLysThrLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu 73
 DB 2114 -----CTGAGATGACGACACCGAGGAGAAATTGATGTGAACGATTGCAACAC 2167
 QY 74 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 93

DB 2168 ACTATCAACGGGATCAATAAAGAGAGACTAGAACGGGACGCTGCTGCCAG 2227
 QY LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrLys 113
 DB 2228 GAGCTGAGAGATACAAAGATCCGTGGATATGAAACACACCTACGCTGCCACT 2287
 QY 114 ValCysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsnGlu 133
 DB 2288 GTGTGCCACCGAATGGC-----AGCTCCCTGACAGCTGACCCAGATCTGACGAATGTG 2341
 QY 134 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalaTyrGluSerTyrSerGlu 153
 DB 2342 ATGGCCACATCCCGAATAATGAAACCTGTATGAGCATGGAGGCTGGAGACAAAG 2401
 QY 154 ValGlyLysGlnLeuArgProLeuTyrGlnGluTyrValValLeuLysAsnGluMetAla 173
 DB 2402 GCGGGAGAGACCATCTCCAGTTTACCGAATAATGAGAACTCATCAACACGAGCTCC 2461
 QY 174 ArgAlaAsnHisTyrGlnAspTyrGlyAspTyrTyrArgGlyAspTyrGluValAsnGly 193
 DB 2462 CGGCTCAATGGCTATGATGATGACAGGGAGCTGTGAGGTCTATGTAGACACACATCC 2521
 QY 194 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGlu 213
 DB 2522 CTGGAG-----CAAGACCTGAGCGGCTCTCCAG 2551
 QY 214 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 233
 DB 2552 GAGCTGACACCACTACTACTCACTGCTTACGCTTACGCTGAGGCGGCGCCCTGACCGTAC 2611
 QY 234 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuGluLysMet 252
 DB 2612 TACGGGGCCACACATCACTGAGGGGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 2671
 QY 253 TrpGlyArgPheThrPheAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 272
 DB 2672 TGGGGGACACCTGCTGCAATCTATGACTTGTGGTGCTGCTGCTGCTGCTGCTGCTGCTG 2731
 QY 273 IleAspValThrAspAlaMetValAspGlnAlaTyrPalaArgIlePheLysGlu 292
 DB 2732 ATGAGACACACAGAGGCTATGCTAAGACAGGCTGAGCCGACGAGAGATGTTAAGAG 2791
 QY 293 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTyrGluAsn 312
 DB 2792 GCTGATGATTTCTTCACTCCCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2851
 QY 313 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTyrAsp 332
 DB 2852 TCGATGCTGAGAGACCAACGAGCGGGAGGTGTCTGTCACAGCTGCGCTGGGAGC 2911
 QY 333 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAspPhe 351
 DB 2912 TTCTACACAGCGAAGAGACTTCGGATCAAGACAGTCCACACCGTCAACTTGGAGAGCCTG 2971
 QY 352 LeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAlaTyrAlaGlnPro 371
 DB 2972 GTGGTGGCCACACCAAGATGGCCACATCCAGTATTCATGACGATCAAGACACTTACT 3031
 QY 372 PheLeuLeuArgAsnGlyValAsnGluGlyPheHisGluValAlaGlyGluIleMetSer 391
 DB 3032 GTGGCTTGGAGGAGGAGTCCCAACCGGCTTCAATGAGGCAATGGGAGGAGCTGTAGCC 3091
 QY 392 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 411
 DB 3092 CTCTAGATGTACGCCACCAAGACCTGACAGATCTCAACCTGCTAGAGAGAGGTGGC 3151
 QY 412 GluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 431
 DB 3152 ACGGAC---GAGCATGATCAACTTCTGATGAAGAGGCGCTTGCACAGATGCGCTTT 3208
 QY 432 LeuProPheThrTyrMetLeuGluLysTyrArgTyrPheMetValPheLysGlyGluIlePro 451

Db 3209 ATCCCTTCAGCTACCTGTCGATCATGCGCGTGGAGGATATTGATGAAGCATCAC 3268
QY 452 LysAspGlnTrpMetLysTrpTrpLysLysAspGlnLysValValGlu 471
Db 3269 AAGGAACTATACCAAGAGAGGTGGAGCTCAGAGTGAAGTACCAAGAGGCTCTGCC 3328
QY 472 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 491
Db 3329 CCAGTGGCCAGGACATCAAGGTGACTTTCACCCAGGCGCCAGTTCACATTCCTCTAGC 3388
QY 492 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnLysLeu 511
Db 3389 GTCCCTTACATCAGGTACTTGTCTCAGCTTCATCATCAGTTCACAGTTCACAGGACACTG 3448
QY 512 CysGlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 531
Db 3449 TCCCAAGCAGCTGGCCACAGGCGCCCTGCACAAAGTGTGACATCTACCACTCCAGAGAG 3508
QY 532 AlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrpThrLeuAla 551
Db 3509 GCCGGGCAAGCGCTGGCGACCGCATGAAGCTGGGCTTCAGTAGGCGGTGGCGGAAAGCC 3568
QY 552 LeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
Db 3569 ATGCAGCTGATCACGGGCGACGCCCAACATGAGCGCCCTCGGCCATGTGAGCTACTTCAAG 3628
QY 572 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 588
Db 3629 CCGCTGCTGAGACTGGCTCGCCAGAGAAAGAGCTGATGGGAGAGAGCTGGGCTGGCCG 3688
QY 589 SerThrAspTrpSerProTyrAlaAspGlnSer 599
Db 3689 CAGTACAACTGAGACCGCAACTCCGCTCGCTCA 3721

Search completed: May 26, 2003, 18:14:45
Job time : 494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 18:03:19 ; Search time 90 Seconds
(without alignments)
2453.414 Million cells updates/sec

Title: US-09-978-385-2_COPY_19_738
Perfect score: 3869
Sequence: 1 STIEBQAKFLDKFNHREAD.....DNSLEFLGIQPTLPPNP 720

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xlh
-O-/cgn2.1/USPTO.spool/US09978385/rnat.21052003.154420.19853/app.query.fasta_1.903
-DB-Issued Patents.NA -OPMT-fastp -SUFFIX-rni -MINMATCH-0.1 -LOOPT-0
-LOOPT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-20
-MODE-LOCAL -OUTFMT-pio -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09978385.ecgn_1.1.40.etrnat.21052003.154420.19853 -NCP-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAEXT-0.5 -DELOP-6 -DELEXT-7

Database: Issued Patents.NA.*

- 1: /cgn2.6/prodata/1/ina/5A.COMB.seq.*
- 2: /cgn2.6/prodata/1/ina/5B.COMB.seq.*
- 3: /cgn2.6/prodata/1/ina/5A.COMB.seq.*
- 4: /cgn2.6/prodata/1/ina/5B.COMB.seq.*
- 5: /cgn2.6/prodata/1/ina/5A.COMB.seq.*
- 6: /cgn2.6/prodata/1/ina/5B.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3869	100.0	2415	4	US-08-989-299-3
2	3869	100.0	3396	4	US-08-989-299-1
3	3214	83.1	2350	4	US-09-280-116-40
4	1337	34.6	2478	1	US-08-481-626-1
5	1337	34.6	4020	4	US-09-050-159-130
6	1337	34.6	4024	4	US-09-162-484-18
7	1310	33.9	3942	4	US-09-162-484-19
8	721	18.6	2082	4	US-09-440-325A-2
9	387	10.0	467	1	US-08-664-596B-14
10	272	7.0	848	4	US-08-905-223-27
11	272	7.0	848	4	US-09-247-155-27
12	133.5	3.5	192	1	US-08-157-171-6

13	123	3.2	8700	2	US-08-392-625-16	Sequence 15, Appl
14	123	3.2	8700	2	US-08-466-961A-16	Sequence 16, Appl
15	121	3.1	789	4	US-09-280-116-114	Sequence 14, Appl
16	120.5	3.1	11384	4	US-08-961-527-45	Sequence 45, Appl
17	120	3.1	144	4	US-08-157-171-4	Sequence 4, Appl
18	120	3.1	11303	4	US-08-961-527-115	Sequence 115, Appl
19	119	3.1	6700	2	US-08-645-193B-118	Sequence 18, Appl
20	116	3.0	2052	3	US-08-630-916A-45	Sequence 16, Appl
21	116	3.0	2052	3	US-09-657-481A-3	Sequence 3, Appl
22	113.5	2.9	1998	3	US-08-844-059-1	Sequence 1, Appl
23	113.5	2.9	1998	3	US-09-431-202-1	Sequence 4, Appl
24	112	2.9	1839	3	US-09-120-365-4	Sequence 4, Appl
25	112	2.9	1839	4	US-09-515-039-4	Sequence 4, Appl
26	110	2.8	3936	4	US-08-961-527-202	Sequence 202, App
27	109	2.8	11309	4	US-08-961-527-108	Sequence 108, App
28	108.5	2.8	9510	4	US-09-453-702B-256	Sequence 256, App
29	107.5	2.8	6506	4	US-09-453-702B-1	Sequence 1, Appl
30	107	2.8	3425	4	US-09-107-149-16	Sequence 16, Appl
31	105.5	2.7	37948	4	US-09-251-645-11	Sequence 11, Appl
32	104.5	2.7	3883	1	US-08-468-036-33	Sequence 33, Appl
33	104.5	2.7	3883	2	US-08-376-843-33	Sequence 3, Appl
34	104.5	2.7	3884	4	US-09-541-782-3	Sequence 3, Appl
35	104.5	2.7	3884	4	US-09-723-820-3	Sequence 3, Appl
36	104	2.7	3672	1	US-08-491-357-1	Sequence 1, Appl
37	104	2.7	3672	3	US-08-968-633-1	Sequence 1, Appl
38	104	2.7	3672	3	US-09-196-466-1	Sequence 1, Appl
39	104	2.7	3672	5	PCT-US96-10823-1	Sequence 1, Appl
40	101.5	2.6	1683	4	US-09-134-001C-316	Sequence 316, App
41	101.5	2.6	1944	2	US-08-844-056-1	Sequence 1, Appl
42	101.5	2.6	3258	4	US-09-595-424-7	Sequence 7, Appl
43	101.5	2.6	21706	4	US-08-961-527-36	Sequence 36, Appl
44	99.5	2.6	2793	4	US-08-953-040-1	Sequence 1, Appl
45	99.5	2.6	10690	4	US-08-961-527-93	Sequence 93, Appl

ALIGNMENTS

RESULT 1
US-08-989-299-3
; Sequence 3, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOGY
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 317
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2415 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-989-299-3

```

Alignment Scores:

Pred. No.:	0	length:	2415
Score:	3869.00	Matches:	720
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-978-385-2_COPY_19_738 (1-720) x US-08-989-299-3 (1-2415)

[illegible]

Db	895	GACACAGGCCCTGGGATGTCACAGAAATATTCAGAGAGGCCGGAATAATGTCCTTTATCTGTT	954
QY	301	GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320
Db	955	GGCTCTCCAAATATGATCTCTCAGAGATTCGGGAAATTTCCATGCTTAACGACCCACAGAAAT	1014
QY	321	ValGlnLysAlaValAlaCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTyrLe	340
Db	1015	GTTTCAGAAAGACAGTCTGCCATGCCACAGCTTGGAGCTGGGGAAAGGGGCACTTCAGAGATC	1074
QY	341	LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis	360
Db	1075	CTTAATGTGCACAAAGGTGCACATGAGCAGCTTCCTGACAGCTCATCATGATGATGGGCAT	1134
QY	361	IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380
Db	1135	ATCCAGTATGATATGGCATATCTGCACACACCTTTTGTCTTAAGAAATGGACCTAATGAA	1194
QY	381	GlyPheHisGluAlaValAlaGlyGluIleMetSerLeuSerAlaAlaThrProCysHisLeu	400
Db	1195	GGATTCACATGAAGCTGTGGGAAATCATGTCATCTTCGACGCCACACCTAGACATTTA	1254
QY	401	LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnDhe	420
Db	1255	AAATCCATTTGGTCTTCTGTCACCCCATTTTCAAGAAAGCAATGAAGAAACAGAAATTAACCTTC	1314
QY	421	LeuLeuLysGluAlaLeuThrTleValGlyThrLeuProPheThrTyrMetLeuGluLys	440
Db	1315	CTGCTCCAAACAGACACTCAGCATTTTGTGGACTCTGCCATTTTACTTAATGTTAAGAGAA	1374
QY	441	TTrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrp	460
Db	1375	TGGAGTGGATGGTCTTTAAAGGGGAAATTTCCCAAGACAGATGATGAAAGTGGTGG	1434
QY	461	GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCys	480
Db	1435	GAGATGAAGCCGAGATATGTTGGGGTGGGAACTGTGCCCATGATGAAGAAACAATACGTT	1494
QY	481	AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg	500
Db	1495	GACCCCGCATCTCTGTCCATGTTCTTAATGATTATTCATTCATTCATTCATTCATTCATTC	1554
QY	501	ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro	520
Db	1555	ACCCTTACCCATTTCCAGTTTCAAAACACACTTGTCAAGACAGCTTAACATGAAGGCCCT	1614
QY	521	LeuHisLysCysAspIleSerAsnSerThrGluIleArgLysLysLeuPheAsnMetLeu	540
Db	1615	CTGGCAACAATGTGACATCTCAAACTCTACAGAACTGGACAGCAAGAACTTTCAATATGCTG	1674
QY	541	ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsn	560
Db	1675	AGGCTTGGAAATTCGAACCCCTGGACCTAGCATTTGGAAGATGGTGTGGAGCAAAAGAAC	1734
QY	561	MetAsnValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGln	580
Db	1735	ATGATATGAAGGCCACTGCTCACTACTTGTAGCCCTTATTACTGGCTGAAGACCGAG	1794
QY	581	AsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProTyrAlaAspGlnSerTle	600
Db	1795	AACAGAGATTCCTTTGTGGAGTGGAGTACCGCAGCTGAGCTCATATGACAGCAAAAGCATC	1854
QY	601	LysValAlaGlyIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn	620
Db	1855	AAAGTGAAGATTAAGCCTTAATACCTTGTGACATTAAGCATATGAAATGAGACACAT	1914
QY	621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysAla	640
Db	1915	GAAATGTACCTGTTCCGATCATCTGTTCATATGCTATGAGGCGAGTACTTTTAAAGATA	1974
QY	641	LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg	660

Db 1975 AAAATCAGATGATCTTTTGGGAGGAGATGTCGAGTGGCTAATTTGAACCAAGA 2034
 QY 115SerPheasnphepValThrAlaProLysAsnValSerAspIleIleProAlaThr 680
 Db 2035 ATCTCTTAAATTTCTTTGTCATGCACTCAAAAATGTCGTATATCTTCTAAGCT 2094
 QY 681 GUAUAGLULysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLysAsn 700
 Db 2095 GAAGTGAAGAAAGGCATCAGAGATGTCGGGAGCCGTATCAATGATGCTTCCGTGAAT 2154
 QY 701 AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2155 GACAACAGCCTAGATTCTTGGGAGATACACCAACTTGGACCTCTTAACCGCCCT 2214

RESULT 2

US-08-989-299-1

Sequence 1, Application US/08989299

Patent No. 6194556

GENERAL INFORMATION:

APPLICANT: Acton, Susan L.

APPLICANT: Robinson, Keith E.

TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG

TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES:

ADDRESSEE: POLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,299

FILING DATE: 11-DEC-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Arnold E. Beth

REGISTRATION NUMBER: 35,430

REFERENCE/DOCKET NUMBER: MIA-025.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3396 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 82..2496

US-08-989-299-1

Alignment Scores:

Pred. No.: 0 Length: 3396
 Score: 3869.00 Matches: 720
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-08-989-299-1 (1-3396)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 136 TCCACCATTTGAGGAGCAAGCCAAAGACATTTTGGACAGATTAAACACAGAGCCGAGAGAC 195

QY 21 LeuPheTyGInSerSerLeuAlaSerTrpAsnTyAsnThrAsnIleThrGluGluAsn 40
 Db 196 CTGTTCTATCAAAAGTCTCTGCTTCTTGAAATTTAAACCAATATATACGAGAGAT 255
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 Db 256 GTCCAAACATGAAATTAATGCTGGGACAAAGTGTGCTTTTAAAGCAACGTCCCA 315
 QY 61 LeuAlaGlnMetTyProLeuGlnIleGlnAsnLeuThrValLysLysLeuGlnLeuGln 80
 Db 316 CTGGCCAAATATATCCACTACAAAGAAATTCAGAAATTCACAGTCAACCTCAGCTGAG 375
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 376 GCTCTTCAGCAAAAGGGCTTCAGTGTCTGACAGAGCAAGAGCAAGCGGTGACACA 435
 QY 101 IleLeuAsnThrMetSerThrIleTySerThrGlyLysValLysAsnProAspAsnPro 120
 Db 436 ATTTCTAAATACAAATGACACCACTACACTACCTGCAAAAGTTTGAACCCAGATATCTCA 495
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTy 140
 Db 496 CAAGATGCTTATTTACTTGAACCAAGTTGAATGAATATGCAACAGTTATGACTAC 555
 QY 141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 Db 556 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGCTGGGCAAGCAGCTGAGCCA 615
 QY 161 LeuTyGluGluTyValValLeuLysAsnGluMetAlaArgAlaAsnHisTyTrsLys 180
 Db 616 TTATATGAAGAGATATGCTGCTTGAATAATGAGATGCGCAAGCAATCATATATGAGAC 675
 QY 181 TyGlyAspTyTrpArgLysAspTyGluValAsnGlyValAspGlyTyAspTySer 200
 Db 676 TATGGGATATATGAGAGAGACATGATGAATATGGGGATGAGATGATGACTACAGC 735
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisTrpPheGlnIleLysProLeuTyGlu 220
 Db 736 CGCGGCCAGTGTATTAAGATGTGAACATACCTTTGAAGAAATTAACCATTAATATGAA 795
 QY 221 HisLeuHisAlaTyValArgAlaLysLeuMetAsnAlaTyProSerTyTrsIleSerPro 240
 Db 796 CATCTTCATGCGCATGTGAGGGGCAAAAGTTGATGAATACCTTCATATCATGTCGA 855
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu 260
 Db 856 ATTGATGCCCTCCCTGCTCATTTGCTTGGATATGTGGGCTAGATTTTGGCAAAATCTG 915
 QY 261 TyrSerLeuThrValProPheGlyLysProAsnIleAspValThrAspAlaMetVal 280
 Db 916 TACTCTTTCAGAGTTCCCTTTGGACAGAAACCAACATGAGTTACTGATGCAATGGTG 975
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGlyAspPheValSerVal 300
 Db 976 GACCAAGCCCTGGGATGACAGAGAAATATCAAGAGCCGAGAAAGTTCTTGTATCTGTT 1035
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 Db 1036 GGTCTTCTCTAATATGACTCAAGATTCCTGGGAAATTCATGCTAACGAGCAAGCAAT 1095
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 Db 1096 GTTCGAAACAGATGCTGCAATCCACAGCTTGGGACCTGGGAAAGGCCAATTCAGAGATC 1155
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
 Db 1156 CTATATGCGCAAAAGGACAAATGAGAGCACTTCTGACAGCTCATATGAGATGGGAGAT 1215
 QY 361 IleGlnTyAspMetAlaTyAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 Db 1216 ATCCAGATATATATGATATGCTGCAACCTTTCTGTAAAGAAATGAGCTAATATGAA 1275
 QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400

```

Db      1276 GGATTCAGTGAAGCGTGGGGAATTCATGCTCTTGTGAGCCACACCTAACCATTTA 1335
Oy      401  LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAspPhe 420
Db      1336 AATCCATTGGCTCTCTGTCACCCGATTTTCAGAGACATGAAACGAAATTAACCTTC 1395
Oy      421  LeuLeuGlyGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGluLys 440
Db      1396 CTGCCTCAACACAGCCTCCACATGTTGGGACTGCTGCATTTACTTACATGTATGAGAG 1455
Oy      441  TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnIlePheLysGlyTrp 460
Db      1456 TGGAGGTGATGGCTTTTAAAGGGGAAATCCCAAGACCGAGGATGAAAGAGTGGTGG 1515
Oy      461  GluMetLysArgGluIleValGlyValGlyIleProValProHisAspGluThrTyrCys 480
Db      1516 GAGATGAGAGCAGAGATGTTGGGGTGGTGGACCTGGCCCATGATGAAACATCTGT 1575
Oy      481  AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db      1576 GACCCCGCATCTGCTTCATGTTCTAAATGATTAATCAATTCATTCATTCATTAACCAAG 1635
Oy      501  ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleHisGlyIlePro 520
Db      1636 ACCCTTACCAATTCAGATTTCAAGAGACCTTTGTACAGCAGCTAAACATGAAAGCCCT 1695
Oy      521  LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeu 540
Db      1696 CTGCACAAATGTGACATCTCAAACTACAGAGACCTGACAGAACTGTTCAATATCTGT 1755
Oy      541  ArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluLysValGlyAlaLysAsn 560
Db      1756 AGCCTTGGAAATATGAAACCTCGGACCTTACCATTTGAAATGTTGAGAGCAAAAC 1815
Oy      561  MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrIlePheLysAspGln 580
Db      1816 ATGATGTAAAGGCCAGCTCACTACTTTGAGCCCTTATTACTTGCTGGAAGACAG 1875
Oy      581  AsnLysAsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIle 600
Db      1876 AACAGAAATCTTTTGGGAGTGGAGTACGACGAGGAGTCCATATGACAGCAAAAGCATC 1935
Oy      601  LysValArgIleSerLeuLysSerAlaLeuGlyAspIleValAlaTyrGluTyrPheAsnAsp 620
Db      1936 AAGATGAGCATTAAGCTTAATACGCTCTTGGAGATTAAGCATATGATGAAACGACAT 1995
Oy      621  GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuVal 640
Db      1996 GAATGTACTGTTCCGATCATCTGTGATATGATATGAGGACAGTACTTTTAAAGTA 2055
Oy      641  LysAsnGlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArg 660
Db      2056 AAAAATCAGATGATCTTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGA 2115
Oy      661  IleSerPheAsnPhePheValIleAlaProLysAsnValSerAspIleIleProArgTyr 680
Db      2116 ATCTCCCTTAATTTCTTGTGCTACGACCTTAATAAATGTCTATATTCATCTTAAGACT 2175
Oy      681  GluValGlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn 700
Db      2176 GAAGTTGAAAGGCCATCAGATGTCGCCGAGCCGATCAATGATGCTTCCGCTGAAAT 2235
Oy      701  AspAsnSerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db      2236 GACACAGCCTAGAGTTTCTGGGAGATACACCAACACTTGGACCTCTTAACCGCCCTCT 2295

```

```

RESULT 3
US-09-280-116-40/c
; Sequence 40. Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Ketch E.

```

```

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280, 116A
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 40
; LENGTH: 2350
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: angiotensin-converting enzyme
; US-09-280-116-40

Alignment Scores:
Pred. No.: 0 Length: 2250
Score: 3214.00 Matches: 706
Percent Similarity: 92.55% Conservative: 2
Best Local Similarity: 92.29% Mismatches: 12
Query Match: 83.07% Indels: 45
DB: Gaps: 2

US-09-978-385-2_COPY_19_738 (1-720) x US-09-280-116-40 (1-2350)
Oy      1  SerThrIleGluGlnAlaLysThr-PheLeuAspLysPheAsnHisGlu-AlaGlu 20
Db      2250 TCCACCATGAGAGAACAGCCCAAGACATTTTGGACAGATTAAACGAAAGCCGAAG 2191
Oy      20  spleuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThr-AsnIleThrGlu 39
Db      2190 ACTGTCTTATCAAAAGTTACTGCTCTTGGAAATTAACCCCAATTAATCTGAAAG 2131
Oy      40  AsnValGlnAsnMetAsnAsnAlaGlyAsp-LysTrpSer-AlaPheLeuGlnGln 59
Db      2130 AATGTCCAAACATGATATATCTGGCGCAAAATGCTGCTTTTAAAGACAGT 2071
Oy      59  erThrLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGln 79
Db      2070 CCACACTTCCCAAAATGTATCCACTACAAAGAAATTCAGATCTCAGCTTCATC 2011
Oy      79  euGlnAlaLeuGlnGlnAsnGlnSerSerValLeuSerGluAspLysSerLysArgLeu 99
Db      2010 TCGAGCTCTTCAACAAATGAGTGTCTCAGCTCTCTGAGAACAGCAAGCAACGCTTGA 1951
Oy      99  snThrIleLeuAsnThrMetSerThrIle-TyrSerThrGly-LysValCysAsnProAs 118
Db      1950 ACACAAATCTTAATTAATCAATGAGCACCACATCATCACTACTGGAAGAAAGTTGTAACCCAGA 1891
Oy      118  pAsnProGlnGlnCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerIle 138
Db      1890 TAAATCCCAAGAAATCTTATTAATCTTGAACCAAGTTTGAATGAATGCAAAACAGTTT 1831
Oy      138  uAspTyrAsnGlnAspGluLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLe 158
Db      1830 AGACTACAAATGAGAGCTCTGGGCTTGGGAAAGCTGGGAGATCTGAGGTGGCAAGCACT 1771
Oy      158  uArgProLeuTyrGlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyr 178
Db      1770 GAGCCCAATTAATGAAGATGATGCTGCTTGAATAATGATGCAAGACAGCAATTCATTA 1711
Oy      178  rGluAspTyrGlyAspTyrTrpArgGlyAspTyrGlu-ValAsnGlyVal-AspGlyTyr 197
Db      1710 TGAGAGCTATGGGATTTATGGAGAGAGACTATGAAATGAATGGGCTTAAGATGCTAT 1651
Oy      198  AspTyr-SerArgGlyGlnLeuIleGluAspValGlnHisThr-PheGlnGluIleLysP 217
Db      1650 GACTACATAGCCCGGCGGCGGATGATGAAAGATGTGACACATACCTTTGAGAGATTAAC 1591
Oy      217  rLeuTyrGlnHisLeuHisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSer 237
Db      1590 CATATATGAACATCTTCATGCTATATGAGGAGCAAGTTGATGAATGCCATCTTCCT 1531
Oy      237  yrIleSerProIleGlyLysLeuProAlaHisLeu-LeuGlyAspMetTrpGly-Argph 256

```

Db 1530 ATATCATGCAATTTGGATGCTCCCTGCTCATTTCCGTTGGATATGGGGTTGATT 1471
 Qy 256 eTrp-ThrasnLeuTySerLeuThrVal-ProPheGlyGln-LysProAsnIle-Asp 275
 Db 1470 TTGGGCAAAATCTGTACTCTTTGACAGTGTCCCTTGGACAGAAACCAACATAGATG 1411
 Qy 275 alThrAspAla-MetValAspGlnAlaTrpAspAla-GlnArgIlePheLysGlnAla-G 294
 Db 1410 TTACTGATCCCAATGGTGGACCAAGCCCTGGATGACCAAGAAATATTCAGAGAGCCCG 1351
 Qy 294 LysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrp-GluAsnSer 313
 Db 1350 AGAAGTCTTGTGATCTGTGGCTCTTCAATATGACTCAAGATTCCTGGGAAATTC 1291
 Qy 314 MetLeuThr-AspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 333
 Db 1290 ATGTAAAGGAGGACCCAGAAATGTTCAGAAAGCAAGTGTGCCATCCACAGCTTGGAGCT 1231
 Qy 333 u-GlyLysGlyAsp-PheArgIle-LeuMetCysThrLysValThrMetAspPhe 352
 Db 1230 GGGGGAGAGGCGAGCTTACAGATCTTAAATGTGCACAAAGTGACAAATGGAGCACTTCT 1171
 Qy 352 uThrAlaHisHisGlyMetGlyHisIleGlnTyraPheMetAlaTyrAlaIleGlnPro 372
 Db 1170 GACAGCTCATCATGAGTGGGCAATCCAGTATGATGCAATATGCTGACACACCTTT 1111
 Qy 372 eLeuLeuArgAsnGlyValAsnGlnGlyPheHisGlnAlaValGlyGlnIleMetSer 392
 Db 1110 TCTGGTAAAGAAATGGAGCTAATGAGATTCATGAAAGCTTGGGAAATCATGTCAC 1051
 Qy 392 uSerAlaIleThrPro-LysHisLeuLysSerIleGlyLeuLeu--SerProAspPhe 411
 Db 1050 TTTCGACGACCAACCTTAAACATTTAAATCCCATGCTGTTTGTGACCCCAATTTTC 991
 Qy 411 GlnGluAspAsnGlnThrGlnIleAsnPheLeuLysGln-AlaLeuThrIleValGly 430
 Db 990 AAGAGACATGAAACGAAATTAACCTCTGCTCAAAACACACACAGATGTTGG 931
 Qy 431 Thr-LeuProPheThrTyMetLeuGlnLysTrpArgTrpMetValPheLysGlyGln 450
 Db 930 ACTCTGCTCATTTACTACTACATGTTAGAGAGTGGAGTGTCTTAAAGGAAAT 871
 Qy 450 eProLysAspGlnTrpMetLys-LysTrpTrpGlnMetLysArgGlnIleValGlyVal 470
 Db 870 TCCCAAGACCAAGTGGATGAAACAAAGTGTGGAGATGAGCAAGATGTTGGGTG 811
 Qy 470 al---GluProVal-ProHisAspGlnThrTyraPheAspProAlaSerLeuPheHisVal 488
 Db 810 TTGGGAACCTTGTGCCCCCATGATGAAACATCTGATGACCCGCACTCTGTTCCATGTT 751
 Qy 489 SerAsnAspTySer-PheIleArgTyTrpArgThrLeuTyr-GlnPhe-GlnPhe 507
 Db 750 TCTATGATTTACTCATTTACTGATTTACACAGAACCCCTTACCATTTCCAGTT 631
 Qy 508 GlnGluAlaLeuLys-GlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIle 527
 Db 631 CAAGAACCACTTGTCCACAGCTAAACATGTAAGGCCCTCTCACAAATGTGACATCTC 631
 Qy 527 r-AsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGln 547
 Db 630 GTAACTCTACAGAGCTGGACAGAACTGTTCATATGCTGAGGCTTGGAAATTCAGAAC 571
 Qy 547 rTrpThrLeuAlaLeuGlnValAsnValAlaGlyAlaLys-AsnMetAsnValArgProLeu 566
 Db 570 CCTGGACCTAGCATTTGAAATGTTGTAGAGCAAAAGAAACATGATGAAGCCACTG 511
 Qy 567 LeuAsn-TyrPheGln-ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 586
 Db 510 CTCAACCTACTTGAAGCCCTTAACTTAACTGCTGAAGACCAAGAAATTTCTTTG 451
 Qy 586 alGlyTrpSerThrAspTrpSerProTyrAlaAspGln-SerIleLysValArgIle-Se 605

Db 450 TGGATGAGTACCGACCTGGAATCCATATGACAGCCAAAGCATCAAAATGAGATAAG 391
 Qy 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGlnTrpAsnAspAsnGlnMetTyLeuPhe 625
 Db 390 CCTAAATATAGCTCTTGGAGATTAAGCATATGATGAAGCAAGCATGAAATGTACTCTT 331
 Qy 625 eArgSerValAlaTyrAlaMetArgGlnTyrrPheLeuLysValLysAsnGln-Met 645
 Db 330 CCGATCATCTGTTGATATGATATGAGGACGACACTTTTAAAGTAAAGAAATCACACATGA 271
 Qy 645 leuPheGlyGlnGlu-AspValArgValAlaAsnLeuLysProArgIleSerPheAsn 664
 Db 270 TTCTTTTGGGAGAGGATGTGCGAGTGTAAATTTGAAACCAAGATCTCTTTAT 211
 Qy 665 PhePhe-ValThrAlaProLysAsnValSerSerPheIleLeuProArgThrGlnValGly 684
 Db 210 TTCTTTGGTCACTGCCCTTAAATATGTGTCTATATCATCTCTAATACTGAAGTTGAAA 151
 Qy 684 salalIeArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer 704
 Db 150 GCCATCAGAGATGTCGCCGAGCCGATCAATGATGCTTCCGTGAATGACACAGCT 91
 Qy 704 uGlnPheLeuGlyIleGlnProThrLeuGlnProAsnGlnProPro 720
 Db 90 AGAGTCTTGGGATACAGCCAAACCTTGACCTCTTAAACAGCCCT 42
 RESULT 4
 US-08-481-626-1
 ; Sequence 1, Application US/08481626
 ; Patent No. 5801040
 ; GENERAL INFORMATION:
 ; APPLICANT: Soudrier, Florent
 ; APPLICANT: Albenc-Gelas, Francois
 ; APPLICANT: Hubert, Christine
 ; APPLICANT: Corvol, Pierre
 ; TITLE OF INVENTION: Nucleic Acid Coding for the Human
 ; TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
 ; TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flinagan, Henderson, Farbow, Garrett &
 ; ADDRESSEE: Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/481,626
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/656,183
 ; FILING DATE: 04-MAR-1991
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 89-09062
 ; FILING DATE: 05-JUL-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meyers, Kenneth J.
 ; REGISTRATION NUMBER: 25,146
 ; REFERENCE/DOCKET NUMBER: 04958-0006-02000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4000
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2478 base pairs

Db 2522 CTGGAG-----CAAGACTGGAGCGGCTCTTCAG 2531

[illegible]

```

Db      3629 CCGCTGCTGAGTGGCTCCGACGAGAGAGAGTGCATGGAGGAGAGAGCTGGCTGGCGG 3688
QY      589 SerThrAspTTPSerProTyrAlaAspGlnSer 599
Db      3689 CAGTACAACTGGACCGCACTCCGCTGCATCA 3721

RESULT 6
US-09-162-484-18
Sequence 18, Application US/09162484
Patent No. 6248724
GENERAL INFORMATION:
APPLICANT: Phillips, M. Ian
APPLICANT: Mohuczy, Dagmara
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
FILE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME MRNA AND METHODS OF USE
FILE REFERENCE: UFLA:087/DEL087P
CURRENT APPLICATION NUMBER: US/09/162,484
CURRENT FILING DATE: 1998-09-25
EARLIER APPLICATION NUMBER: 60/059,661
EARLIER FILING DATE: 1997-09-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 4024
TYPE: DNA
ORGANISM: Homo sapiens
US-09-162-484-18

Alignment Scores:
Pred. No.: 1,256-149 Length: 4024
Score: 1337.00 Matches: 255
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 34.56% Indels: 34
DB: 4 Gaps: 9

US-09-978-385-2_COPY_19_738 (1-720) x US-09-162-484-18 (1-4024)
QY      2 ThrileugluginlalaThyrPheleuaspPheasnHlsgluAlaAspLeu 21
Db      1952 ACTGATGAGGCGTACGAGCAGCAAGATTGTGGAGGAATGTGACCGGACATCCAGGTGGTG 2011
QY      22 PheTyrGlnSerSerleuAlaSerTyrPasnTyrPasnThrAsnIleThrGlnGlu 39
Db      2012 TGGACGAGATGCGCGAGCCAACTGACATCAACCAATCCACACGAGACGAGCAGC 2071
QY      40 -----AsnValGlnAsnMetAsnAlaGlyAspTyrSerAla 53
Db      2072 AAGATTCTGCTGACAGAAACATGCAATACCCAAACACAC----- 2113
QY      54 PheleuYsgluginSerThrleuAlaGlnMetTyrProleugluginIleGlnAsnLeu 73
Db      2114 -----CTGAGTACGCGACCCAGCCAGGAAAGTTGATGTGACCACTTCAGAACACC 2167
QY      74 ThrValYsgluginleuGlnAlaLeuGlnGlnAsnGlySerSerValleuSerGluAsp 93
Db      2168 ACTATCAAGCGGATCAATAAGAGGTTCAGACCTAGAACCGGACGCTCCCTGCCAG 2227
QY      94 LysSerTyrArgleuAsnThrIleleuAsnThrMetSerThrIleTyrSerThrGlyLys 113
Db      2228 GAGCTGGAGAGTACACAAAGATCTCTGTGATATGAAACACCACTTACAGCTGGCCACT 2287
QY      114 ValCysAsnProleuAsnProGlnGluCysleuLeuLeuGluProGlyLeuAsnGluIle 133
Db      2288 GTGCTCCACCGCAATAGC-----AGCTGCGTGCAGCTCGAGCCAGATCTGACGAATGTG 2341
QY      134 MetAlaAsnSerleuAspTyrAsnGluArgleuThrAlaTrrpGluSerTrrpArgSerGlu 153
Db      2342 ATGGCGACATCCCGGAAATATGAAGACCTGTATGTGGCATGGAGGGCTGGCGAGACAG 2401
QY      154 ValGlyLysGlnleuArgProleuTyrGlnGluTyrValleuLysAsnGlnMetVal 173

```

```

Db      2402 GCGGAGAGAGCCATCTCCAGTTTACCCGAAATACCTGGAATCATCAACGAGGCTGC 2461
QY      174 ArgAlaAsnHlsglYsglYsglYsglYsglYsglYsglYsglYsglYsglYsglYsglYsgl 193
Db      2462 GCGCTCAAGGCTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2521
QY      194 ValAspLysTyrAspTyrSerArgGlyGlnleuIleGluAspValGlnIleThrPheGlu 213
Db      2522 CTGGAG-----CAAGACCTGGAGGCGCTCTCCAG 2551
QY      214 GluIleLysProleuTyrGlnHlsglAlaTyrValArgAlaLysleuMetAsnAla 233
Db      2552 GAGCTGAGCCACTTCACTCAACCTGATGCTAGCTGAGCGCGCGCGCTGACCGCTCAC 2611
QY      234 Tyr-----ProserTyrIleSerProIleGlyCysleuProAlaHlsglLeuGlnAspMet 252
Db      2612 TACGGGGCCAGACATCAACCTGGAGGGGCGCATCTCTGCTACCTGCTGGGGAATG 2671
QY      253 TrrpGlyArgPheThrPasnleuTyrSerleuThrValProPheGlyGlnLysProAsn 272
Db      2672 TGGCGGAGACTGTGTCAAACATGTGACTGTGTGCTGCTCCCTCCAGCCCGCTCG 2731
QY      273 IleAspValThrAspAlaMetValAspGlnAlaTrrpAspAlaGlnArgIlePheLysGlu 292
Db      2732 ATGGACACCCACAGGCTATGCTTAAAGCAGGCGTGGAGCCGCCAGGAGGATGTTTAAAGAG 2791
QY      293 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnIleYhrTrpGluAsn 312
Db      2792 GCTGATGATTTCTTCACTCCCTGCGGGGCTGCTCCGCTGCTCTGATGTTTGAAACAG 2851
QY      313 SerMetleuThrAspProGlyAsnValGlnLysAlaValCysHlsglProThrAlaTrrpAsp 332
Db      2852 TCGATGTGGAGAAAGCAACCGAGGCGGAGAGTGTGTCCACCCCTCGCGCTGGGAGC 2911
QY      333 LeuGlyLysGly-----AspPheArgIleLeuMetCysThrLysValThrMetAspPhe 351
Db      2912 TTCTACACGCGCAGAGACTTCGCGATCAAGCACTGACACCACTGACCTGGAGACCTCG 2971
QY      352 LeuThrAlaHlsglMetGlyHlsglIleGlnTyrAspMetAlaTrrpAlaIleGlnPro 371
Db      2972 GTGGTGGCCACACGAAATGGGCGCATTCATGATTCGATGCAAGACTTACT 3031
QY      372 PheleuLeuArgAsnGlyAlaAsnGluGlyPheHlsglAlaValGlyLylleuMetSer 391
Db      3032 GTGGCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3091
QY      392 LeuSerAlaAlaThrProLysHlsglLysSerIleGlyLeuLeuSerProAspPheGln 411
Db      3092 CTCTCAAGTCTTACGCGCAAGCAACCTGACAGCTGACACTGCTGACAGTGAAGGCTGGCC 3151
QY      412 GluAspAsnGluThrGluIleAsnPheleuLeuLysGlnAlaLeuThrIleValGlyThr 431
Db      3152 AGCGAC-----GAGCATGACATCAACTTCTGATGAAATGATGCGCTTACAAAGATCGCCTT 3208
QY      432 LeuProPheThrTyrMetleuGluLysTrrpArgTrrpMetValPheLysGlyLylleuPro 451
Db      3209 ATCCCTCTCAAGCTACCTCGTGCATGAGTGGCGTGGAGGAGGATTTATGAAAGCATCACCC 3268
QY      452 LysAspGlnThrMetLysTrrpTrrpGluMetLysArgGluIleValGlyValGlu 471
Db      3269 AAGGAGAACTATTAACAGAGAGGTGTGAGGCTGAGGCTGAAGTACCAAGGCGCTGCGCC 3328
QY      472 ProValProHlsglAspLysThrTyrCysAspProAlaSerleuPheHlsglSerAsnAsp 491
Db      3329 CCAATGCCCGCAGGACTCAAGTGTGACTTGTGACCGGCGGCAAGTTCACATCTCTCTTAC 3388
QY      492 TyrSerPheIleArgTyrTrrpArgTrrpThrLeuTyrGlnPheGlnHlsglAlaLeu 511
Db      3389 GTGCTTACATCAAGGATCTTGTACCTTATATCACTTCACTTCAAGTTCACAGAGGAGCTG 3448
QY      512 CysGlnAlaAlaLysHlsglGluProleuHlsglCysAspIleSerAsnSerThrGlu 531
Db      3449 TGGCAGGACACTGCGCACACGCGGCGCCCTGCAAAATGTGACATCTTACCAAGTCCAAAGAG 3508

```

QY 532 AlAGlyInLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProThrPheLeuAla 551
DB 3509 GCGGGGAGCGCGTGGGAGCGCCATCAACGCTGAGTGCAGTAGCGCGCGGAGGAGC 3568
QY 552 LeuGlnAsnValAlaGlyAlaLysAsnMetAsnValArgProLeuAsnTyrPheGlu 571
DB 3569 ATGCAGCTGATACAGGGCCAGCCCAACATACAGCGCCCTGCGCATGTGGACCTACTTCAG 3628
QY 572 ProLeuPheThrTyrLeuLysAspGlnAsnLys-----AsnSerPheValGlyTyr--- 588
DB 3629 CCGCTGCTGAGTGGCTCCGACGAGAACGAGCTGATGGGAGAAAGCTGGCTGGCG 3688
QY 589 SerThrAspTyrProTyrAlaAspGlnSer 599
DB 3689 CAGTACAACTGGAGCGGCACTCGCTGCTCA 3721

RESULT 7
US-09-162-484-19
Sequence 19, Application US/09162484
Patent No. 6248724
GENERAL INFORMATION:
APPLICANT: Phillips, M. Ian
APPLICANT: Mohuczy, Dagnara
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME mRNA AND METHODS OF USE
FILE REFERENCE: DPLA:087/DPLA087P
CURRENT APPLICATION NUMBER: US/09/162,484
EARLIER FILING DATE: 1998-09-25
EARLIER APPLICATION NUMBER: 60/059,661
NUMBER OF SEQ. ID NOS.: 19
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 19
LENGTH: 3942
TYPE: DNA
ORGANISM: Rattus norvegicus
US-09-162-484-19

Alignment Scores:
Pred. No.: 2,1e-146 Length: 3942
Score: 1310.00 Matches: 251
Percent Similarity: 61.20% Conservative: 115
Best Local Similarity: 41.97% Mismatches: 214
Query Match: 33.86% Indels: 18
Gaps: 7

US-09-978-385-2_COPY_19_738 (1-720) x US-09-162-484-19 (1-3942)

QY 2 ThrTllegLugLugLlnAlaLysThrPheLeuAspLysPheAsnHisGluAspLeu 21
DB 1948 ACTGATGAAACCCAGGCTAACAGCTTCGTGAGAGTATGACCGGACGCCAAGGTGTG 2007
QY 22 PheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnLleThrGlnGluAsnVal 41
DB 2008 TGAAGCAATACAGCAGGCGCACTGATTAACACCACTTTCATAGAGGGCAGC 2067
QY 42 GlnAsnMetAsnAlaGlyAspLysTyrPheAlaPheLeuLysGlnGlnSerThrLeu 61
DB 2068 AATACCTCTGCTTCAAGAAACAAAGAGTGTCCAACTTGAATATGCGACTCTGG 2127
QY 62 AlAGlnMetTyrProLeuGlnGluLlnGlnAsnLeuThrValLysLeuGlnLeuGlnAla 81
DB 2128 GCCAAGACATTTACGTGACCACTTCCAGAACTCACTCAAGCGGATCATMAAGAG 2187
QY 82 LeuGlnGlnAsnLysSerValLeuSerGluAspLysSerLysArgLeuAsnThrIle 101
DB 2188 GTTCAGACAGCTGACCGGAGTGTGCTCCCAAGAGATTAGAAAGATACCAACGATC 2247
QY 102 LeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGln 121
DB 2248 CTGCTAGACATGTGAGACGACTTACAGTGTAGCCAAATGTTTGCATACAAATATGCG----- 2301

QY 122 GluCysLeuLeuLeuGluProGlyLeuAsnGluLlnIleMetAlaAsnSerLeuAspTyrAsn 141
DB 2302 ACTGTCTCTCACTGAGCGCTGATCTGACAAATATAAAGGCACAGCTCCCGGAATATACGA 2361
QY 142 GluArgLeuThrAlaThrGlnSerTyrPheSerGluValGlyLysGlnLeuArgProLeu 161
DB 2362 GAATTCCTTGTGGTGTGGAGAGCTGGCAGACAAAGTGGGAGAGCCATCTTCCCTTT 2421
QY 162 TyrGlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyr 181
DB 2422 TTCCCAAACTAGCTGGACTTCTTCACAAAGATGCCAGGCTACAGGCTACTCTGATGCA 2481
QY 182 GlyAspTyrTyrPheGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArg 201
DB 2482 GGGGATTCCTGGAGATCTCATATGATGATCGGATGATGAG----- 2523
QY 202 GlyGlnLeuLlnGlnAspValGlnLlnThrPheGlnGluLlnLysProLeuTyrGlnHis 221
DB 2524 -----CAAGACCTGGAAAACATATACAGAGCTGACCGCTCTACTCTGAC 2571
QY 222 LeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSer---TyrIleSerPro 240
DB 2572 CTGCATGCTTATGTGGCGCGCTCCCTGCACCGCATTTATGGGTGAGTATACATCACTG 2631
QY 241 TllegLysCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260
DB 2632 GATGGTCCCATCTGCTCCACCTGCTAGGAGAACTGGGACACAGACTGTGTCACAACTC 2691
QY 261 TyrSerLeuThrValProPheGlyLlnLysProAsnIleAspValThrAspAlaMetVal 280
DB 2692 TATGACTTGTGGCACCCCTCCCTCCGCCCCCACTATAGATGACCGAGGAGCTGATTA 2751
QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerVal 300
DB 2752 AAGCAGGATGACACACCAAGAGATTTTAAGAGAGTGCACAAATTTTAACTCTCCCTG 2811
QY 301 GlyLeuProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsn 320
DB 2812 GGGCTGTACCTGCTGCCCCCTGAGTGTGGAACAAATGTTGAGAACCAACGAT 2871
QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPheLeuGlyLysGly---AspPheArg 339
DB 2872 GGGAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2931
QY 340 IleLeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGlnMetGly 359
DB 2932 ATCAAGCAATGTAACCTGTGTGAACATGGAGATGTGTGATAGCCACACAGAAATGGCC 2991
QY 360 HisTllegLysAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsn 379
DB 2992 CACATCCAGTATTTATCATGACAGTACAAAGACTTGTGCTGCTGCTGCTGCTGCTGCTG 3051
QY 380 GlnGlyPheHisGlnAlaValGlyLlnLysLeuMetSerLeuSerAlaAlaThrProLysHis 399
DB 3052 CCGGCTTTTCAAGAGGCTATTGAGATGTGTTGGCTGTCTGTCTGTCTGTCTGTCTGTCTGT 3111
QY 400 LeuLysSerTllegLysLeuLeuSerProAspPheGlnGlnLysAsnGluThrGlnLeuAsn 419
DB 3112 CTACACAGCTTCACACCTGCTGACAGTGTAG---GGCAGTGGCTGACAGCATGATCATC 3168
QY 420 PheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlu 439
DB 3169 TTTCTAATGAAGATGGCCCTTGACAAAGTCCCTTCAATCCCTTCAAGCTACCTCATGAC 3228
QY 440 LysTyrPheTyrPheValPheLysGlyLlnLysLeuProLysAspGlnTyrMetLysTyr 459
DB 3229 CAGTGGCCCTGGAGGCTTTTACGAGCATACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3288
QY 460 TrpGlnMetLysArgGluLlnLysValGlyValValGlnProValProHisAspGluTyr 479
DB 3289 TGAAGTCTACAGCTAAGTACAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3348
QY 480 CysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThr 499

Oy	GIUGLUGlnAlaLeuSthrPhrLeuAspLysPheAsnHISGLuAlaGlnAspLeuPheYtr	23
	::: ::: ::: ::: :::	
Db	157 GAAACGARGARCNAAARHTHTTTTCNCARTTTTATGATCACCACGCGAGCTGNTATNAAY	211
Oy	24 GlnSerSerLeuAlaSerTrpAsnTYrAsnThrAsnIleThrGIUGlnAsnValGlnAsn	43
	::: ::: ::: ::: ::: :::	
Db	217 AARTTYATGARGCNCNCTGGAAATYATGTACNACNAAYATGACMGNAABAAYCARGARGAR	276
Oy	44 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGIUGlnSerThrLeuAlaGln	63
	::: ::: ::: ::: ::: :::	
Db	277 ATGATGARGAAYATGARGMGN--MSNCARTTYATGATHTATTTGGNACNCARCACNCAAY	333
Oy	64 MetCysProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln	83
	::: ::: ::: ::: ::: ::: ::: :::	
Db	334 YLTNTTYAARSTNACNCARTTYYARAGATCCNGATGTMNAAYGNAATGYTWTMSNAARTTCAR	392

QY	84	GLNANGLSYSESRVALLSESRGLSPYSSSESLYSAGVLEASNTLHLEUASN	103
DB	394	AAATNHAAGVARGCGCNYTWMSNAARATGARTYNNMCGARTAAAGARTYNTNCCN	453
QY	104	ThrmcSerThrIletyrSerThrglyysValcysasnProaspasnproglnclycs	123
DB	454	NNNTNGARATGACATAWSSNATGGCCNARGINTGYTNAAYGAAGGCCN-----TGY	507
QY	124	LeuleuleuGluProGlyLeuasnGluIleMetAlaasnSerLeuaspTyrAsnGluArg	143
DB	508	YTMSNTYNGAR---WSNGARTYNGARSTMTATGGCCNACNWSMGNGAlAAGARGARTTN	564
QY	144	LeuTrpAlaTrpGluSerTyrArgSerGluValGlyLysGlnLeuArgProLeuTyrGlu	163
DB	565	YTNTGGCGNTGGCARGNTGGCARGAYCGNGTNGMNCARATHHTGYACNACNTTGGAR	624
QY	164	GluTyrValValLeuLeuLysAsnGluMetAlaArgAlaasnHis---TyrGluAspTyrGly	182
DB	635	CAYTAATGNGARYYTNWSNAAYAAACCCNCCARYYTNAAYGGGTNNNAARAAVATGGGN	684
QY	183	AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly	202
DB	685	GCNTYNTGGCAVWSNAARTAYGARNSNACNTYNGAR-----	723
QY	203	GlnLeuIleGluAspValGlnHisThrpheGluGluIleLysProLeuTyrGlnHisLeu	222
DB	724	-----CARGAYTNGARGMGTNTTTCARGARYYTNMNCNCCNYTATYATNAAYCCN	774
QY	223	HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyr---ProSerTyrLisSerProIle	241
DB	775	CATACNTATYGTNNMGMGNCNTYTNCAVMGNCATYTAGGCCNCGARTYNTATGATYTNMGN	834
QY	242	GlyCysLeuProAlaHisLeuLeuGlyAsp---MetTrpGlyArgPheTrpThrAsnLeu	260
DB	835	GGNCNCAHCCNCCNCAATYTYTNGNCNARAAVACNTYNGCCNCAWSMTGGTNAAYATH	894
QY	261	TyrSerLeuThrValProPheGlyGlnLysProasnIleaspValThrAspAlaMetAla	280
DB	885	YTNGAYCCNGTNTYTNCCNTTYYTNARARARATGCCNARGAIGTNACNAARATATGARGAR	954
QY	281	AspGlnAlaTrpAspAlaGlnArg---IlePheLysGluAlaGluLysPhePheValSer	299
DB	955	GTNCARCAVYTGGAARCCNARGARARATYNTATGTNGARGARGCNGARACNTTYYTTCNTAY	1012
QY	300	ValGly-----LeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAsp	317
DB	1015	YTNGNNTYNGCNYTNCCNCCNCCNCCNCCNMSNTTYYTGAAARATYNTATGTNTATGNGN	1074
QY	318	ProGlyAsnValGlnLysAlaValAlaCysHisProThrAlaTrpAspLeuGlyLys---Gly	336
DB	1075	CNMCNCAVYGGMGMGARGGTNGARTGycATATWMSNGCNTGGAATYYTATCARAGATGAY	1133
QY	337	AspPheArgGlyLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHis	356
DB	1135	GAYTTYMNAHNAARARATGYCGNARGRTNCNCCNCGARGAYCCNYTWMSNATHTTTCAY	1194
QY	357	GluMetGlyHisIleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsn	376
DB	1195	CARATGGCNCATYYTCARTATYYTNTNCARATAYARAAYTYTNWSNATATHTTYGNAACN	1254
QY	377	GlyAlaAsnGluGlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThr	396
DB	1255	GGNGCNAAVCCNGCNTTGTGARGARGCNGTNGMNSGNTNATHACNYTNMNSCNMSNSN	1314
QY	397	ProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThr	416
DB	1315	CAYARCAKYTYNTYNAATATRGNTYNTYTNMSN-----YTYTNACARAY	1355
QY	417	GluIleAsnPheLeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyr	436
DB	1360	GARGTNAAYTYTYTNTATGCATAATGCCNTTNGARARARATHGCTYATATGCCMTTGGCNTAY	1419
QY	437	MetLeuGluLysTrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMet	456

848
60
23
37
6
3
27 (1-848)

```

...:::|||||
lysseraLaLeugLyasplys 613

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

phgglvgllngllasnva}Arar 653

GAATTTCCCATGTCCTA 250

ValThrAlaProLysAsnVal 673

GGTTACAGACCTTCAAA--- 307

```
ArgmetSerArgSerArgIle 693
```

AGGAAAGAACAGAACCGAATC 367

[illegible]

the Angiotensin I Conver

2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645
2646
2647
2648
2649
2650
2651
2652
2653
2654
2655
2656
2657
2658
2659
2660
2661
2662
2663
2664
2665
2666
2667
2668
2669
2670
2671
2672
2673
2674
2675
2676
2677
2678
2679
2680
2681
26


```

: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue, NW
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/466,961A
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/392,625
: FILING DATE: 22-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/876,791
: FILING DATE: 30-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/784,234
: FILING DATE: 31-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652.0980004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-466-961A-16

Alignment Scores:
Pred. No.: 0.00266 Length: 8700
Score: 123.00 Matches: 149
Percent Similarity: 32.36% Conservative: 116
Best Local Similarity: 18.19% Mismatches: 250
Query Match: 3.18% Indels: 304
DB: 2 Caps: 39

US-09-978-385-2-COPY_19_738 (1-720) x US-08-466-961A-16 (1-8700)
QY 2 ThrllegluglgnlnalalystrPherleuasplyspheasnHlsglu----- 17
| | | | | : : : | | | | | : : : | |
Db 2619 ACTATTAGGAAATTTATTCACGAATTTTTTATGTGATATAATTGATTTGGACAACACTAGTAAT 2678
| | | | | - - - - - AlaGluspleu- - - - - PheTyrglnSerSerleuaIleSertPrasn 31
QY 18 -----AlaGluspleu-----PheTyrglnSerSerleuaIleSertPrasn 31
| | | | | : : : | | | | | : : : | |
Db 2679 TTAAAGCAATTCCTCTCAGATATTAATGATTTGGCTATCCCAAAAAAGACAGTATAGT 2738
| | | | | - - - - - TyrAsnThrAsnIleThrGlugluasValglnasmEtaAsnAlaglyAsplysrTrp 51
QY 32 TyrAsnThrAsnIleThrGlugluasValglnasmEtaAsnAlaglyAsplysrTrp 51
| | | | | : : : | | | | | : : : | |
Db 2739 TTTTCTAATAACATTT----- 2753
| | | | | : : : | | | | | : : : | |
QY 52 serAlaphelaulysgluginserThrleualaglnmetyrProleuglnglwllecin 71
| | | | | : : : | | | | | : : : | |
Db 2754 ---GCATTTTAAAAGAAAG----- 2771
| | | | | : : : | | | | | : : : | |
QY 72 AsnleuthrValLysleuglnleuglnaleuglnglnaanglySerSerValleuSer 91
| | | | | : : : | | | | | : : : | |
Db 2772 -----TATTTGCTGTCATTCAAAATTAACAGCATATTTGAATAATACA 2813
| | | | | : : : | | | | | : : : | |
QY 92 GluasplysserysarglueasnthrIleleuasnthrMetSerThrlleTySerThr 111
| | | | | : : : | | | | | : : : | |

```

[illegible]

1 CLASSIFICATION: 424
 2 PRIOR APPLICATION DATA:
 3 APPLICATION NUMBER:
 4 FILING DATE:
 5 ATTORNEY/AGENT INFORMATION:
 6 NAME: Brooks, A. Anders
 7 REGISTRATION NUMBER: 36,373
 8 REFERENCE/DOCKET NUMBER: PB340P1
 9 TELECOMMUNICATION INFORMATION:
 10 TELEPHONE: (301) 309-8504
 11 TELEFAX: (301) 309-8512
 12 INFORMATION FOR SEQ ID NO: 45:
 13 SEQUENCE CHARACTERISTICS:
 14 LENGTH: 11384 base pairs
 15 TYPE: nucleic acid
 16 STRANDEDNESS: double
 17 TOPOLOGY: linear
 18
 19 JS-08-961-527-45

Alignment Scores:	
Pred. No.:	0.00837
Score:	120.50
Percent Similarity:	31.28%
Best Local Similarity:	17.98%
Query Match:	3.11%
Ob:	4
Length:	1138
Matches:	146
Conservative:	108
Mismatches:	238
Indels:	321
Gaps:	45

US-09-978-385-2_COPY_19_738 (1-720) x US-08-961-527-45 (1-11384)

QY	5	GLGCLINLALYSTTIRPHELEUSAPRLYSRHEASNIISGLVLALEIUSPREUPHE	YrGln 24
Db	11096	GAAGCATTAATAAACCTTATG	: : : : :
QY	25	SeSerLeuAlaSer	-----ACGGGTGAAATAATTTTATCTCCAA 11055
Db	11051	CATTATATGAGACACATAGGAGAACTAATATGAGAGCAGGTATACCTCCGTGT	-----29
QY	30	TRPASNTRYASN	-----ThraSnIleThrGluGln 39
Db	10991	TGGGCACCACTATGCTCAGCGCTGTTCACCTGGTGTGATTTTACCAACTGGATTGAAAT	: : : : : 109312
QY	40	ASNVALIGLINSMETASNANALAGLYASPLRYTSPSERALAPHELEUYSGLGINSER	59
Db	10931	CAGATTCCATGTGTAAGAAATGATTTTGGGCTGTGGAACTCTT	: : : : : ACC 10884
QY	60	ThLeuAlaGln	-----MetTyrProLeuGlnIuIleGlnAsnLeuThrValYls 76
Db	10883	AATATGCTCAAGAGGGCATATTTAC	-----10857
QY	77	LeuGlnLeuGlnAlaLeuGlnGlnIuAsnGlySerSerValIleuSerGlnLysSerLys	96
Db	10856	AAATATCATGTGCACACCTCAAAATGTCATCAACTATGAG	: : : : : 10815
QY	97	ArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsn	116
Db	10814	ATGACCTTGGCTGCAGATAGAGCT	-----10785
QY	117	ProAspAsnProGlnGluCysLeuLeuGlnIuProGlyLeuAsnGlnIleMetAlaAsn	136
Db	10784	CGTCCAGGAACAAGGGCATTCGTAACA	: : : : : 10755
QY	137	SerLeuAspTyrArgnGluArgLeuTyr	-----AlaTrpGluSerThrArgSerGlu 153
Db	10757	GAGCTTCCTCGTGAAGAAATGAAGATGAGACTTGGCTGCACGAAGAAACGT	: : : : : 10704
QY	154	ValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMetAla	173
Db	10703	TGGGGCTTTGAAGAGCGTCCCTGC	-----10688
QY	174	ArgAlaAsnIleTyrGlnAspTyrArgLysSerTyrTrpArgGlyAspTyrGluValAsnGly	193
Db	10679	AATTTATATGAGTTCACGCTGATCTGGAAA	-----ACAAAT 10641

QY	194	ValaspGlyTyrAspTyrSerAlaGlyGlnLeuIleGlnAspValGlnHisThrPheGlu	213
Db	10640	TCGTATGGCAGCTCTTATAGTATTTGGCCACCTAAAGATGAACTCTTCCTTATCTCGTT	10581
QY	214	GlnIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla	233
Db	10580	GAAATGGAAC-----TTACTACATATATGAG-----	10557
QY	234	TyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrp	253
Db	10556	-----TTTATGCC-----TTATGCCCCATCTTTGGCTTGAGTGTGG	10518
QY	254	GlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlnLysPro-----	271
Db	10517	GGGTATCAGCTTATGGGTACTTCCGTTTAGACATGCTTATGGCCGACCAAGAGAGATT	10458
QY	272	-----AsnIleAspValThrAspAlaMetValAsp	281
Db	10457	CAGATTTTGTTCGAGAGAGTGTCTACACCAATAAATATGGGGT-----ATTGGGAC	10407
QY	282	GlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGly	301
Db	10406	-----TGGGTACCACTGCTACT	10392
QY	302	LeuProAsn-----MetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGly	319
Db	10391	TTACCACATCAACGATGATGCTCTTA-GCCTATTATGATGGACACCGACTTTGGAATACCA	10333
QY	320	AsnAlaGlnLysAlaValCysHisProThrAlaTrp-----AspLeuGly	334
Db	10332	GACCATTAATAGGCT-----CATACCACTGTTGGGGTCCCTTAATTTTGACCTTGGGA	10279
QY	335	LysGlyAspPheArg-----IleLeuMetCysThrLysValThrMetAspAspPhe--	351
Db	10278	AAAATGTAAGTCCAGTCCCTCTTAATTTCTTGATTAAGATGATGATGATGATGATGAT	10219
QY	352	-----LeuThrAlaHisHisGlnMetGlyHisIleGlnTyrAspMetAla	366
Db	10218	TTGGATGCTATTCGTGTGATGCTGTGTTAGCAACATCTCTATTTTGACATGATGATGCT	10159
QY	367	TyrAlaIleAlaGlnProPheLeu-----LeuArgAsnGlyAlaAsn-----GlnGlyPhe	382
Db	10158	-----CCATGGACACTAATTAAGATGGCGGAATGTCACATGATGAAGTTAT	10111
QY	383	His-----GlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeu	400
Db	10110	TATTTCTCTTACGGCTTGATGGATGAGTTATTAAGTTAGAA-----	10072
QY	401	LysSerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGlnIleAsnPhe	420
Db	10071	-----TATCCAGATGTCATGATGATATTCCTCGTTTACGAAAGAAAGTTCCGTGCG	10030
QY	421	LeuLeuLysGlnAlaLeuThrIleValaGlyThrLeuProPheThrTyrMetLeuGlnLys	440
Db	10029	ATCAAGATTTACGGGAATGAAGAAGATATGGTGCCTAGATTGACTAC-----AAA	9979
QY	441	TrpArg-----TrpMet-----ValPheLysGlyGlnIlePro-----	451
Db	9978	TGGAACTGGGCTGGATGAATGATATTCCTCGTTTACGAAAGAAAGATCCGATCTATCGT	9919
QY	452	-----LysAspGlnTrpMet	456
Db	9918	AAATATGACTTTAACTCGTGTGACTTTCAGCTTATGATGATTTTCAAGAGAAATATATCC	9859
QY	457	LysLysTrpTrpArgIleMetLysArgGlnIleValaGlyValaGlnProValProHisAsp	476
Db	9858	TTG-----CCATTCCGACGAT	9841
QY	477	GlnThrTyrCysAspProAlaSerLeuPheHisValSer-----AsnAsp	491
Db	9840	GAAGTGGTTGACGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	9781
QY	492	TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeu	511

```

Db      9780 TTCGAGGCTG-----CGCAATCTCTATACGTACCA-----ATT 9745
QY      512 CysGlnAlaAlaLysHis----- 517
Db      9744 TGTACACCTCGTAAGAAATGTGCTTCATGGGTACGGAATACGGTCATTCCTAGAAATGG 9665
QY      518 -----GluGlyProLeu----- 521
Db      9684 AAATCTGAAGACAGTGGTAATGTCTTAACCTAGAAAGCCCAATGAATGTAAGATGAAG 9625
QY      522 -----HisLysCys----- 524
Db      9624 TATTTCGCTTCACCTAACCAAGCTTTTACAAAGATCATCGCTGCTCTGTGGAAATTCAT 9565
QY      525 -----AspIleSerAsnSerThrGluAlaGluLysLeuPheAsn 538
Db      9564 ACCAGCTATGATGATTAATCAATCATGTGCGGTAATTCAGACCAAGAGTTCCTTCC 9505
QY      538 MetLeuArgLeuGlyLysSerGluProThrPheLeuAlaLeuGluAsnValValGlyAla 558
Db      9504 TTTATTCGTAAAGGTAAGGAGAAATGTTAGTCTGATCTTAATATGTTACCTGTT 9445
QY      559 LysAsnMetAsnValArg-----ProLeuLeuAsnTyrPheGluProLeuPhe--- 574
Db      9444 GACCGGAAAGATTTTACATCGACTACCCGTTGACAGAAATTTACGAAAGATGAGAAAT 9385
QY      575 ThrTrpLeuLysAspAsnGlnAsnLysAsnSerPheValGlyTrpSerThrAspTrpSerPro 594
Db      9384 ACTGATGTGGAAGAG-----TGGGAGGCGCTTGGAAAGAA 9349
QY      595 TyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAspLysAla 614
Db      9348 CAT---AAACCAACGCTTCAACGCAAGAGACTA----- 9316
QY      615 TyrGluTrpAsnAspAsnGlnMetTyrLeuPheArgSerSerValAlaTyrAlaMetArg 634
Db      9315 -----TGGAAAGATTATGACGACGACTTAACCTTACCCGCTATGGGACCAAGT 9262
QY      635 GlnTyrPheLeuLys-----ValLysAsnGlnMetIleLeuPheGlyGluGluAspVal 652
Db      9261 GATGCAAAATCAACGCTGCTTGAATCTACTAAACCGTCACAAATAAAACCAAAA 9202
QY      653 ArgValAlaAsnLeuLysProArgLysSerPheAsn 664
Db      9201 GGAGTGAAGAAATGAATAATGTAGCTTGTAT 9166

RESULT 17
US-08-157-171-4
; Sequence 4, Application US/08157171
; Patent No. 5736323
; GENERAL INFORMATION:
; APPLICANT: Soubrier, Florent
; APPLICANT: Hubert, Christine
; APPLICANT: Corvol, Pierre
; TITLE OF INVENTION: Agents and Procedures for the Study of
; TITLE OF INVENTION: The Genetic Polymorphism of the Angiotensin I Converting
; TITLE OF INVENTION: Enzyme
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 90 South 7th Street, 3100 No. 5736323west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,171

```

```

; FILING DATE: 24-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hillson, Randall A.
; REGISTRATION NUMBER: 31,838
; REFERENCE/DOCKET NUMBER: 8076.103USMO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 144 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Exon 17
; US-08-157-171-4

Alignment Scores:
Pred. No.: 5.87e-06 Length: 144
Score: 120.00 Matches: 21
Percent Similarity: 59.57% Conservative: 7
Best Local Similarity: 44.68% Mismatches: 19
Query Match: 3.10% Indels: 0
Gaps: 0

US-09-978-385-2_copy_19_738 (1-720) x US-08-157-171-4 (1-144)

QY      130 LeuAsnGlnIleMetAlaAsnSerIleAspTyrAsnGluArgLeuTrpAlaTrpGluSer 149
Db      3 CTGACGAATGTGATGGCCACATCCCGGAATTAAGAACCTGTATGGGATGGAGGCG 62
QY      150 TrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlnGluTyrValLeuLys 169
Db      63 TGGGAGACAAAGAGGGGGAGAGGACCATCTCCAGTTTACCCGAAATACGTGGAACTCATC 122
QY      170 AsnGluMetAlaArgAlaAsn 176
Db      123 AACGAGGCTGCCGCGCTCAAT 143

RESULT 18
US-08-961-527-115/C
; Sequence 115, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1

```


TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:

LENGTH: 11303 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-115

Alignment Scores:

Pred. No.:	0.0095	Length:	11303
Score:	120.00	Matches:	131
Percent Similarity:	32.36%	Conservative:	101
Best Local Similarity:	18.27%	Mismatches:	187
Query Match:	3.10%	Indels:	298
DB:	4	Gaps:	42

US-09-978-385-2_COPY_19_738 (1-720) x US-08-961-527-115 (1-11303)

QY 22 PheTyrGlnSerLeuAlaSerTrpAsnTyrAsnThr---AsnIleThrGlnGluAsn 40
 DB 2860 TTTTTCATTTGACTATTATTGTCATGATATATACAGATAGAAACATTCGATTCAGATATA 2801
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeu-----LysGlu 57
 DB 2800 ATGAAACACGCTCAAAATG-----GTGCGTCTGCTTTTATGCTACCAATAAAAGGAG 2747
 QY 58 GlnSerThrLeu-----AlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu 73
 DB 2746 GAAACAAATTTGAAAAATTTTAAAGCGAGACTTTTAT-----CGAATCTCTGATTTA 2696
 QY 74 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 93
 DB 2695 ACAGACACAGCTAGTAGATGATGATCTCAAACTGAGAGT-----TTTGAGGAT 2645
 QY 94 LysSerLysArgLeu-----98
 DB 2644 AGAAGCAGACAACTTATCTATATATAGAGATTTGGCCAAATTCGTATGATTTATATCGAG 2585
 QY 99 ---AsnThrIleLeuAsnThrMetSerThrIle-----TyrSerThrGlyLys 113
 DB 2584 ACATCAACAGTACTAGATATATATGACGAGCTATTTTCATCTGCACCTTTAATAAGGAGAA 2525
 QY 114 ValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIle 133
 DB 2524 TTT---AACCCCGCAAAAT-----TTATTTATA-----GAATTA 2495
 QY 134 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGlu 153
 DB 2494 AAGGAAATTTTGGCTATATGAGTAGTCTGCTGATATTTGGTGAATCTATATGAACGTTA 2435
 QY 154 ValGlyLysGlnLeuArgProLeu-----TyrGlnGluTyr 165
 DB 2434 AAGGAA---ACAGCTCGAGCTCTGAGTAAGCAGATGACAGCAATCTCGTTAGAGAGATGGA 2378
 QY 166 ValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrp 185
 DB 2377 ATATATGCGCGAATGCTATAGCTTATTTGGAT-----GATTCCTCAAGGATTTATTTGG 2324
 QY 186 ---ArgGlyAspTyr-----GluValAsnGlyValAsp-----195
 DB 2323 GATGACAAATCTGCTTATATTTGAGAGAGAAATCAAAATTCACCTGATTTCTTTGCCATAT 2264
 QY 196 GLYTyrAspTyrSerArgGlnLeuIleGluAspValGluHisTrpPheGlnGluIle 215
 DB 2263 GGGGATGATTTAT---AGAGGA---GCAATCAGAGATTTTACCAT-----2225
 QY 216 LysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrPro 235
 DB 2224 ---TTGACTGCTGCTCAACACC 2207

QY 236 SerTyrIleSerProIleGlyCysLeuProAlaHisLeuGlyAspMetTrpGlyArg 255
 DB 2206 ---TTGTGCCAAGATATGCTTTAGCAATTTGGGAGTATGGA 2168
 QY 256 PheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsnIleAspVal 275
 DB 2167 TATTTGG-----CCATTACGCTCGGATGAAATCTGATTTTA 2132
 QY 276 ThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLys 295
 DB 2131 ATAGACAGATTT-----GAAACAGAGAAA 2108
 QY 296 PhePheValSerValGlyLeuProAsnMet-----305
 DB 2107 ATTCCATTATCTATCTGCTGTAGATATGATGATGATGATGATGATGATGATGATGATGAT 2048
 QY 306 ---ThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyLysn 320
 DB 2047 TTTGAGAGTGGCTGCGACAGATATAGTGGAAATAGAAACTTATATCAAAATCCAGAACAG 1988
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1987 TTATTCACAACTT---CATGATGAAAGCTTAAACTCTCC-----1949
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHis 360
 DB 1948 ---TTAAATGTCATCTGCTGATGGATTA 1922
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyLysAsnGlu 380
 DB 1921 CGGCGTTATGAAAGAACTTAT-----CCTCAAAATCCGA 1889
 QY 381 GlyPheHisGlnAlaValAlaGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1900 ---LysSerPro 408
 QY 401 LysSerIleGlyLeu-----1889
 DB 1888 AAACGGTGGGCTTAAATGATAGTACAGTAAAGAACTGATTTTGAATTTTAAATCC 1829
 QY 409 AspPheGlnGluAspAsnGluThrGlnIleAsnPheLeuLeuLysGlnAlaLeuThrIle 428
 DB 1828 TCTTTTAAAGGAAAGCTTAAAGTGTTCAT-----1796
 QY 429 ValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArgTrpMetValPheLysGly 448
 DB 1795 ---TATGACACAGAAAAG-----1781
 QY 449 GlnIleProLysAspGlnTrpMetLysLysTrpTrp---GluMetLysArgGluIleVal 467
 DB 1780 ---CAGGAGTAGATTTTGTGGATTTGATGCGCAACAGGACACACAA 1736
 QY 468 GlyValValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHis 487
 DB 1735 GGTATCTGATTCACCTT-----1718
 QY 488 ValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeuTyrGlnPheGlnPhe 507
 DB 1717 ---TGGCTTTAAACATATATACATAT 1694
 QY 508 GlnGlnAlaLeuCysGlnAla-----AlaLysHisGlnGlyPro---520
 DB 1693 CAGGATAGTGTGTAATAATGCAAGAGTGGTTGATTTATCAAGATATGACAGCTCTGCT 1634
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeu-----PheAsn 538
 DB 1633 AGTCACCGCTACCCGCTTGTTTCA-----GGGAGTACTATTTATTTAGTTGCAAT 1583
 QY 539 MetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAla 558
 DB 1582 TCTTTAAAGATTT---CAACCTATTTTACAGGACAGCATCTAATATCGGT---1535
 QY 559 LysAsnMetAsnValArgProLeuLeuAsnTyrTrpPheGluProLeuPheThrProLys 578

Db 1534 -----TANAGTTGG----- 1526
 QY 579 AspGlnAsnLysAsnSerPheValGlyThrSerThrAspTrpSerProTyrAla----- 596
 Db 1525 -----TGGAGTCATGANTGCGTGACATATGTTGGGG 1493
 QY 597 -----AspGlnSerIleLysValArg-----IleSerLeuLysSer 608
 Db 1492 GATTATGACGAGAGCTACAACTGATGCGTACAGTTTGCTTTTGTAGTCCGATTACT 1433
 QY 609 AlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgSerSer 628
 Db 1432 CGATTACATGATGTTAGAGTCTTTTATGCTAAAGAACCTTGCTTTTTCACAGAACCA 1373
 QY 629 ValAlaTyrAlaMetArgIleThrPheLeuLysValLysAsnGluMetIle 645
 Db 1372 ACATCTAAGATTATGAGAAATATC--CTTCGTTGAGACATCAATGATT 1325

RESULT 19

US-08-645-193B-18
 ; Sequence 18, Application US/08645193B
 ; Patent No. 5962253

; GENERAL INFORMATION:
 ; APPLICANT: Kupke, Thomas
 ; APPLICANT: Goltz, Friedrich
 ; APPLICANT: Kempter, Christoph
 ; APPLICANT: Jung, Gunther
 ; TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
 ; TITLE OF INVENTION: Catalyzed by Flavoprotein Epld
 ; NUMBER OF SEQUENCES: 70
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/645,193B
 ; FILING DATE: 13-MAY-1996
 ; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 0652,1540000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ. ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8700 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; US-08-645-193B-18

Alignment Scores:

Pred. No.: 0.00804 Length: 8700
 Score: 119.00 Matches: 150
 Percent Similarity: 32.248 Conservative: 115
 Best Local Similarity: 18.25% Mismatches: 247
 Query Match: 3.08% Indels: 310
 DB: 2 Gaps: 40

US-09-978-385-2_COPY_19_738 (1-720) x US-08-645-193B-18 (1-8700)

QY 2 ThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlu----- 17
 Db 2619 ACTATATGGAATATATCAGCAATTTTATGCAATATATATGATTTGAACAACATGATAAT 2678
 QY 18 -----AlaGluAspLeu-----PheArgIleSerLeuAlaSerTrpAsn 31
 Db 2679 TTAAGCAATTCCTCTCAGATATATATGATTTGGCTATCCAAAAAGACACTTATAGT 2738
 QY 32 TyrAsnThrAsnIleThrGluLysValGlnAsnMetAsnAlaGlyAspLysTrp 51
 Db 2739 TTTCTTAATATACAT----- 2753
 QY 52 SerAlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnGluIleGln 71
 Db 2754 ---CCATTTTAAAGAAAG----- 2771
 QY 72 AsnLeuThrValLysLeuGlnIleGlnAlaLeuGlnGlnAsnGlySerSerValLeuSer 91
 Db 2772 -----TATTTGCTTCGCAATTCMAATTAACAGCCATATTTGAATTAACA 2813
 QY 92 GluAspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThr 111
 Db 2814 GAAACGACGCTTAAATTTTGAAGAAC--AATATATCACTTCTTAAATCAATGCG-- 2867
 QY 112 GlyLysValLysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsn 131
 Db 2868 -----CCTGTTTCACAT 2879
 QY 132 GluIleMetAlaAsnSerLeuAspTyrAsnGluArgLeuThrAlaTrpGluSerTrpArg 151
 Db 2880 GAAATTA-----TATAGGAGATATATATTT----- 2903
 QY 152 SerGluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlu 171
 Db 2904 -----GCAATTCGATTAAGAAAGT--TATGAGATTTTCCCGATCAATCCAAAT 2951
 QY 172 MetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgLysAspTyrGluVal 191
 Db 2952 TTAGGACCTTTTAAATGCGCGTCAACCTTTTGAAGGTTT--ACGGGAAATTTCAATATA 3008
 QY 192 AsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGluHisThr 211
 Db 3009 AAG-----AAAAAATCAATTTAAGAAAGAAATATGTCAT-- 3044
 QY 212 PheGluGluLeuLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMet 231
 Db 3045 -----CATTTACATATATATCAATGATGAATGAATGCTTA 3077
 QY 232 AsnAlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAsp 251
 Db 3078 GAAATTAAGCCCAATTAATGAAGGTCCTTAACCTCAAGAAATGTAAATATTTGCAATAT 3137
 QY 252 MetTrpGlyArgPheTrpThrAsnLeuTyrSerIleLeuThrValProPheGlyGlnLysPro 271
 Db 3138 -----AATGCAATATATATCTGTTTAATTTAATTTACT-----AAAGT 3182
 QY 272 AsnIleAspValThrAspAlaMetValAspGlnAlaThrAspAlaGlnArgIlePheLys 291
 Db 3183 GATATAGATATATTAATGACATTTATTTGAGGCTACATTTAACAACCTTATCATATATCT 3242
 QY 292 Glu-----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGly 308
 Db 3243 GAAAAACATGATTCAGAAATTTGTATGATCTAATTCATTAATGTTTAATTTGAGTTTGA 3302
 QY 309 -----PheTrpGluAsnSerMetLeuThrAspPro 318
 Db 3303 TCTGATATATACAAATTTTAAAGAAATTTTCATTTGAAAAACAAATTTATATACACCT 3362
 QY 319 GlyAsn-----ValGlnLysAla 324
 Db 3363 ATTAAGTGAAGAACGATTCATTAACCTTTTGTCCAGAAATTAATTAATAAAATATAT 3422
 QY 325 ValCysHisProThrAlaThrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThr 344

```

Db      3423 ATTTTAAACGACGCTTGGAAAAATTAATTCAGAAATGTTTCTGAA-----ACT 3473
Qy      345 LysValThrMetAspSerPheLeuThrAlaHisHisGlyHisIleGlnTyrAsp 364
Db      3474 GAAATATGCTTAAATGTTCCGCACTATT---AGAGAAAAATGGCAATATCCAAAAAGAT 3530
Qy      365 MetAlaTyrAla-----
Db      3531 GTATATATGCTTTTGGAGATATATGATGCTATTAATTAATTAATGACAAGCATCTTC 3590
Qy      369 -----AlaGlnProPheLeuLeuArgAsnGlyAla 378
Db      3591 ATATATCTAAAAAGACTAAAAAACAATGGTAGATTCGAATATTAAGAAAGCTTATC 3650
Qy      379 AsnGlnGlyPheHisGlnAlaValGlyIleMetSer----- 391
Db      3651 AATGAATCTATATATAGAGAAATGTAGAAATTTGACCATTAATTAATAAAACTACT 3710
Qy      392 -----LeuSerAlaAlaThrProLys-----HisLeuLys 401
Db      3711 TTTAAAGAACAAATCTTTCATTAATACCTAAATAATAGAACATTCATATATCTTAA 3770
Qy      402 -----SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGlnTyrGluIle 418
Db      3771 GATTGGTTTCAATTCATTAATTAATGATCTTAAACATACCAAGAT----- 3815
Qy      419 AsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeu 438
Db      3816 AATTTTATTCAGAAATATATATATATATATATATATATATATATATATATATATAT 3875
Qy      439 GlnLysThrArgTyrMetValPheLysGlyIleProLysAspGlnTyrMetLysLys 458
Db      3876 AATAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3923
Qy      459 TyrTrpGlnMetLysArgGluIleValGlyValValGluProValProHisAspGluThr 478
Db      3924 ---TTAAGATTATTAAGAGAA-----GATGAAGAT 3950
Qy      479 TyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr 498
Db      3951 TAT-----TTCCAATTAATTAATTCATTAATTAATTAATTAATTAATTAAT 3983
Qy      499 -----ThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCys 512
Db      3984 AAGATTATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4043
Qy      513 GlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSerThrGluAla 532
Db      4044 GAAGTATATAGATATGCTGTCACACAGTAAATGAGATATGAGATTTTATTAATGAT 4103
Qy      533 GlyIleLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnTyrProThrThrLeuAla 552
Db      4104 GATGCTCTATTA-----TCAATTAATTAATTAATTAATTAATTAATTAATTAAT 4154
Qy      553 GluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGluPro 572
Db      4155 GAATTTATGCTTGTCT-----ATATCAATTAATTAATTAATTAATTAATTAAT 4202
Qy      573 LeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTyrPserThrAspTyr 592
Db      4202 ----- 4202
Qy      593 SerProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAsp 612
Db      4203 -----ATTATTAATAAGTAGAGAAAGAAAGAAATTTTATTAATTAATGCG-----GAA 4247
Qy      613 LysAlaTyrGlnTyrPheAsnAspAsnGlnMetLysLeuPheArgSerSerValAlaTyrAla 632
Db      4248 GATTTAATTCCTAGTAAATGAC----- 4268
Qy      633 MetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGluGluAspVal 652

```

```

Db      4269 ATAGAGATAT----- 4280
Qy      653 ArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsn 672
Db      4281 -----AAAAATTATATGCTAAACT-----ACCAATCTCAAAAT 4316
Qy      673 ValSerAspIleIlePro----- 678
Db      4317 GACTATGAATTTTAAAAAAGAAATTCGCAATCTCATGAAATTTTATTAATTAAT 4376
Qy      679 -----ArgThrGluValGlnLysAlaIleArgMetSerArgSerArg 692
Db      4377 AGTATTTAGAAATCTTAAAAAGACACTACAAAAAAGCTTATTAATTAATTAAT 4436
Qy      693 IleAsnAspAlaPhe-----ArgLeuAsnAspAsnSerLeuGluPheLeuGlyIle 709
Db      4437 AATATGCGACGTTTATATACATACGCTGTGTAAT-----AGAAATTCGTAAT 4484
Qy      710 GlnPro 711
Db      4485 AATCCT 4490

RESULT 20
US-08-630-916A-45
Sequence 45, Application US/08630916A
Patent No. 601137
GENERAL INFORMATION:
APPLICANT: Pirozzi, Gregorio
APPLICANT: Kay, Brian K.
APPLICANT: Fowlkes, Dana M.
TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,916A
FILING DATE: 03-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-203
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 896-8864/9741
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2052 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-630-916A-45

Alignment Scores:
Pred. No.: 0.0016 Length: 2052
Score: 116.00 Matches: 141
Percent Similarity: 32.17% Conservative: 107
Best Local Similarity: 18.29% Mismatches: 265
Query Match: 3.00% Gaps: 268
DB: 36

```

US-09-978-385-2_COPY_19_738 (1-720) x US-08-630-916A-45 (1-2052)

OY		12	AspLysPheAsnHisGluValIleGluAspLeuPheThrLeuSerLeuAlaSerTrpAsn	31
Db		152	GAACTGTATTGGAGATCATCTCTCA---TTTGCACCAACTGTAATAAAGCGTCGTCCAGC	208
OY		32	TyrAsnThrAsnIleThrGlnGluAspValIGlnAsnMetAsnAsnAlaGlyAspLysTrp	51
Db		209	GGTAGCTCCAGTAGTGTCGAAGAANAATGCCCTGTCTCCAAATTGGCACT-----	256
OY		52	SerLAlaPheLeuLysGlnGlnInsertThrLeuAlaGlnMetCyrProLeuGlnGlnIle--	70
Db		257	-----AQTACTACTGTGGAGATCCCTCCAGTCCAGTTCAAGAAATACGT	295
OY		71	-----GlnAsnLeuThrValLysLeuGlnLeuGlnIleValLeuGlnGlnInsangly	86
Db		296	ACTTCCTCAGAAAACAATGAATATATTCCTTACCACAGTCCACAATTTGGATCTGAAGCT	355
OY		87	SerSerValLeuSerLysSplysSerLysArgLeuAsnThrIleLeuAsnThrMetSer	106
Db		356	AGAAAGATATTATTAGACGCTCACACCTCTMAAT-----TCTAAGAGTAGT	397
OY		107	ThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnIucyLysLeuLeu	126
Db		398	TCGCGCTTTTGAGACACCAATTAACAACGCCAGATGGGTATGATGATCCTGTACGCGAG	457
OY		127	GluProGlyLeuAsnGlnIleMetAlaAsnSer-----	137
Db		458	CAGTCTGGG-----AATGCCAACACAGAAACCTTGGCATCAGGGTGGCAACA	505
OY		138	-----LeuAspTyrAsnGlnIucylgLeuThrAlaTrp	147
Db		506	AGAAAAAGATCCCTCAGTGAAGAACTTATGTGGACAAATAAATCTCGAACTCCACATG	565
OY		148	GlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlnGlnIucylVal	167
Db		566	GAGAGA-----CCACAACTTTAACTCCGACGGTTGGGAAAGAGAGTTGAT	610
OY		168	LeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGly	187
Db		611	GATCGTAGAAGATTATTATGTGATGTCATTAACCCAGAACAAACATCGGACGGCGCT	670
OY		188	AspTyrGlyL-ValAsnGlyValAspGlyTyrAspTyrSerArgLysIleuIleGlu	206
Db		671	ACCANAGGAATCTGTCCGAATTTTGAACAGTGGCAATCTCACGGGACCAATTTG----	724
OY		207	AspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisIleuHisAlaTyrVal	226
Db		725	-----CAGGGAGCTATGACMACAGTTTAAACACAGATACCTTATTCGGCTCAATGTTA	778
OY		227	ArgAlaLysLeuMetAsnAlaTyrProSerTyrIleserProIlglyCysLeuProAla	246
Db		779	GCTGCAGAA-----AATGACCTTATGAGACCTTGTGCACCA	814
OY		247	HisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValPro	266
Db		814	-----	814
OY		267	PheGlyGlnLysProAsnIleAspValThrAspAlaMet-----	279
Db		815	-----GGCTGGGAAAAAGATGGATTCCAACAGACAGGTTTTACTTTGTGAATCATTAACACA	871
OY		280	-----Valasp	281
Db		872	AAACAAACCCAGTGGAGATCCAAAGACTCAAGCTTACAGAAATGAAGAAMCCCTGCCA	931
OY		282	GlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGly	301
Db		932	GAAAGCGCTGGGAAATTT--AGATATATCTCGTGAAGGTGTGAAGTACTTTGTTGATCATCAAC	988
OY		302	LeuProAsnMetThrGlnGlyPheTrpGluAsnserMetLeuThrAspProGlyLysVal	321
Db		989	ACAAAGACAAACAACA-----TTCAAAGATCTCGGCAATGGG	1024

OY	322	GLNLALALALALysHis-----	ProthAlatrAspLeuGlyLysAspRhe	338
Db	1025	AAATCAATCTGTAACTAAAGGTGGTCCAAATATTCCTTATGCAAGCGGC-----	TTT	107
OY	339	ArgIleLeuMetCysThrLysValThrMetAspAspRheLeuThrAlaHisGluMet	358	
Db	1076	AGG-----	TGGAAGCTT	108
OY	359	GlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnProRheLeuLeuArgAsnGlyAla	378	
Db	1088	GCTCACTCCGTTAT-----		110
OY	379	AsnGlnGlyPheHisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLys	398	
Db	1103	-----	TTGGCAGCTTAATCACTACCTACTAG	112
OY	399	HisLeuLysSerIleGlyLeuLeuSerProAspRheGlnGluAlaAspAsnGluThrGluIle	418	
Db	1130	CATGTAAAG-----ATCAATGTGCCGCGACAGACTTGTGGAGATTCCTTCCAAAGAT	118	
OY	419	AsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProRheThrTyrMetLeu	438	
Db	1187	---ATGGCATTTAAA-----	CCCATGACTGTG	121
OY	439	GluLysTyrArgTyrMetValPheLysGlyGluIleProLysAspGln-----	TrpMet	456
Db	1211	AGAGAGCGCGTTATATGATATATTTAGAGGAGAAGAAGACTTGATATGGTGGCTAAGCG	127	
OY	457	LysLysTyrTrpGlnMetLysArgGluIleValGlyValGlnProVal-----		473
Db	1271	AGAGATGGTTTTTTCTTGCTTACAT-----	GAAGTTTAAACCAATGATAGCTTA	132
OY	474	-----ProHisAspGluThrTyrCys-----	AspProAlaSerLeuHe-----	486
Db	1325	TTTAGTAGTACGGGGAAGAACACTATGTCTGCGAGATTAATCCAGCAATCAACCATTAAT	138	
OY	487	-----HisValSerAsnAspTyrSerPheIle-----	ArgTyrTyrThrArgThrLeuTyr	503
Db	1385	CCAGACCACTTTCTCAATAC-----TTCTGTTTCATGGTGGTTTATATGCCATGGCACTAATTT	144	
OY	504	GlnPheGlnPheGlnGluAlaLeuLysGlnAlaAlaLysHisGluGlyProLeuHisLys	523	
Db	1442	CATGAAAGTTATATGCATACAGCTGGTTCTCTTACCATCTTCAACAGCGTATGTAAAGTAA	1501	
OY	524	-----CysAspIle-----		526
Db	1502	AAACTTACTATTAAGGATTTGAATCTATGTACTGATGAATTTTAACTCCCTTATCTGG	1561	
OY	527	-----SerAsnSerThrGluAlaGlyGlnLysLeuPhe-----	AsnMetLeu	540
Db	1552	ATAGACAGTAAACACATTTGAAGAAATGGCTGTAGAAATGTAATCTTTCTGTGGACATGGAG	1621	
OY	541	ArgLeuGlyLysSerGlnProTyrPheAlaLeuGlnAsnValAlaGlyAlaLysAsn	560	
Db	1622	ATTTTGGGAAAGTACTTCACTACATGACCTGAAGTTG-----	CGAGGTTCAT	1665
OY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGln	580	
Db	1670	ATTCTGGTG-----	ACTGAGAG	1687
OY	581	AsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIle	600	
Db	1688	AACAAGATGATATATTTGGTTTAAATGACAGAAATGG-----		1723
OY	601	LysValArgIleSerLeuLysSerAlaLeuLysAspLysAlaTyrGluTyrPasnAspAsn	620	
Db	1724	-----CGTTTCTTCGAGAGAGTCAAGAACAGACCAAGAGCTTCTCTGATGGCTTTAAT	1777	
OY	621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640	
Db	1778	GAAGTT-----	GTTCCTCTTCAGTGGGTACAGTACTTCGATGAGAAAA	1819

QY 641 LysAsnGlnMetIleLeuPheGlyGluAspValArgValAlaAsnLeuLysProArg 660
Db 1820 GAATTAGAGTTATGTTGTGGCATGCAGAGTTGACTTGGCAGATTGGCAGAAAT 1879
QY 661 IleSerPheAsnPhenPheValThrAlaProLysAsnValSer-----AspIleIle 677
Db 1880 ACTGTTATCGACATTATACAAAGAAACAGCAATCATTTGTTTGGCAGTTTGTG 1939
QY 678 ProArgThrGluValGluLysAlaIleArgMet 688
Db 1940 AAAGAGCAGACATGATGAGTAAGAAATGCGACTA 1972

Search completed: May 26, 2003, 20:22:58
Job time : 150 secs

10

11

12

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 18:06:24 ; Search time 268 Seconds

(without alignments)
3547.511 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738

Perfect score: 3869
Sequence: 1 STRIBQAKFLDKFHEAD.....DNSLEFLGIQPLGPNQPP 720

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=fragnet_p2n.model -DEV=rxh
-Q=/cgn2_1/USFRO.spool/US09978385/runat_21052003_154420_19915/app.query.fasta_1_903
-DB=Published.Applications_NA -QFMT=fastap -SUFFIX=trpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100
-THR.MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09978385 -ECGN_1_1_57 -runat_21052003_154420_19915
-NCPU=6 -ICPU=3 -NO.MMAP -LARGEOUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3869	100.0	3334	9	US-09-978-385-1
2	3869	100.0	3369	9	US-10-158-847-141
3	3865	99.9	3325	9	US-10-114-893-85
4	3833	99.1	2920	9	US-09-969-384-2

5 3833 99.1 2920 9 US-10-158-847-137
6 3785 97.8 2911 9 US-09-969-384-12
7 3720 96.1 3732 9 US-10-028-072-71
8 3720 96.1 3732 9 US-10-121-049-71
9 3720 96.1 3732 9 US-10-123-904-71
10 3720 96.1 3732 9 US-10-140-470-71
11 3720 96.1 3732 9 US-10-175-746-71
12 3720 96.1 3732 9 US-10-176-918-71
13 3720 96.1 3732 9 US-10-176-921-71
14 3720 96.1 3732 9 US-10-137-865-71
15 3720 96.1 3732 9 US-10-140-474-71
16 3720 96.1 3732 9 US-10-142-431-71
17 3720 96.1 3732 9 US-10-143-114-71
18 3720 96.1 3732 9 US-10-140-002-71
19 3720 96.1 3732 9 US-10-142-419-71
20 3720 96.1 3732 9 US-10-123-262-71
21 3720 96.1 3732 9 US-10-142-423-71
22 3720 96.1 3732 9 US-10-121-050-71
23 3720 96.1 3732 9 US-10-143-755-71
24 3720 96.1 3732 9 US-10-143-032-71
25 3720 96.1 3732 9 US-10-123-108-71
26 3720 96.1 3732 9 US-10-123-236-71
27 3720 96.1 3732 9 US-10-123-261-71
28 3720 96.1 3732 9 US-10-140-921-71
29 3720 96.1 3732 9 US-10-140-928-71
30 3720 96.1 3732 9 US-10-121-045-71
31 3720 96.1 3732 9 US-10-123-292-71
32 3720 96.1 3732 9 US-10-123-903-71
33 3720 96.1 3732 9 US-10-124-819-71
34 3720 96.1 3732 9 US-10-124-822-71
35 3720 96.1 3732 9 US-10-140-925-71
36 3720 96.1 3732 9 US-10-160-498-71
37 3720 96.1 3732 9 US-10-121-041-71
38 3720 96.1 3732 9 US-10-121-043-71
39 3720 96.1 3732 9 US-10-121-047-71
40 3720 96.1 3732 9 US-10-123-215-71
41 3720 96.1 3732 9 US-10-123-902-71
42 3720 96.1 3732 9 US-10-123-908-71
43 3720 96.1 3732 9 US-10-123-909-71
44 3720 96.1 3732 9 US-10-123-910-71
45 3720 96.1 3732 9 US-10-124-813-71

ALIGNMENTS

RESULT 1
US-09-978-385-1
: Sequence 1, Application US/09978385
: Patient No. US20020177211A1
: GENERAL INFORMATION:
: APPLICANT: Piddington, Christopher S.
: APPLICANT: Petrie, Charles
: APPLICANT: Shoemaker, Kimberly E.
: APPLICANT: Bishop, Paul D.
: TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
: FILE REFERENCE: 99-24C1
: CURRENT APPLICATION NUMBER: US/09/978, 385
: CURRENT FILING DATE: 2001-10-16
: PRIOR APPLICATION NUMBER: 60/133, 952
: PRIOR FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: 60/151, 181
: PRIOR FILING DATE: 1999-08-27
: PRIOR APPLICATION NUMBER: 09/563, 516
: PRIOR FILING DATE: 2000-05-03
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 3334
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (35)...(2449)

Db 2069 AACTCTTTAATTTCTTCTTCTACCTGCACCTAAAAATGTCTGATATCATTTCTTAGAACT 2128
Qy 681 GIUVALGULYSALALALARGMETSERARGSERARGILEASNPALAPHEARGLEUASN 700
Db 2129 GAAGTTGAAAGAGCCATCGAGATGTCGCCGAGCCGATCAATGATGCTTTCGCTCGAAT 2188
Qy 701 ASPASNSEULEUGLUPHELEUGLYILEGINPROTHLEUGLYPROPROASNGINPROPRO 720
Db 2189 GACACAGCCCTAGAGTTCTGCGGAGATACAGCCACACTTGACCTTCATACAGCCCTCT 2248

RESULT 2
US-10-158-847-141
; Sequence 141, Application US/10158847
; Publication No. US20030091557A1
; GENERAL INFORMATION:
; APPLICANT: Tom Parry et al.
; TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
; FILE REFERENCE: PP557
; CURRENT APPLICATION NUMBER: US/10/158,847
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,004
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-158-847-141

Alignment Scores:
Pred. No.: 0 Length: 3396
Score: 3869.00 Matches: 720
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-10-158-847-141 (1-3396).

Qy 1 SerThrIleGluGlnAlaIleAlaIleThrPheLeuAspIlePheAsnIleGluAlaIleAsp 20
Db 136 TCCACCTTGAAG 195
Qy 21 LeuPheTyrGlnSerSerLeuAlaIleSerThrPheAsnIleThrPheAsnIleThrGluGluAsn 40
Db 196 CTGTTCTATCAAAAGTTCCTGCTTCTGGAATTAATACCAATATTAATCTGAAAGAGAAAT 255
Qy 41 ValGlnAsnMetAsnAsnAlaIleAspIleThrPheLeuAspIleThrPheAsnIleThrGluGluAsn 60
Db 256 GTCCAAACAGAAATTAATGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315
Qy 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValIleLeuGlnIleLeuGln 80
Db 316 CTTCGCCAAATGATTCACATCAAGAAATTCAGAAATTCACAGTTCAGCTCAG 375
Qy 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspIleThrPheAsnIleThrGluGluAsn 100
Db 376 GCTCTCAGCAAAATGGGCTTCAGTCTCTCAGCAAAATGGGCTTCAGTCTCAGTCTCAG 435
Qy 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyValIleCysAsnProAspAspPro 120
Db 436 ATCTTAATATCAATAGAGACCATCTACATCTACATCTGAAAGTTTAAACCCAGATTAATCCA 495
Qy 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
Db 496 CAGAAATGCTTATTAATCTTGAACAGGTTTGAATTAATGCAAAAGTTTGAATTAATGCA 555
Qy 141 AsnGluArgLeuThrPheIleThrPheGluSerThrPheGluValIleGlyGlnIleLeuArgPro 160
Db 556 AATGAGAGGCTCTGCTGGGAG 615
Qy 161 LeuTyrGlnGluIleTyrValIleLeuIleCysAsnGlnMetAlaArgAlaAsnIleThrGluAsp 180

Db 616 TTATATGAGAGAGATGTTGGTCTTGAATAATAGATGGCAGAGAGAGAGAGAGAGAGAGAGAG 675
Qy 181 TyrGluAspTyrThrPheGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
Db 676 TATGGGATTTTGGAG 735
Qy 201 ArgGlyGlnLeuIleGluAspValGluIleThrPheGluGluIleThrPheGluGluIleThr 220
Db 736 CGCGGCGAGTTGATTAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 795
Qy 221 HisLeuHisAlaTyrValArgAlaIleLeuMetAsnAlaTyrProSerTyrIleSerPro 240
Db 796 CATCTTCTGCTTATGATGAG 855
Qy 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrThrAsnLeu 260
Db 856 ATTGATGCTCCCTCCCTGCTCATTTGGTGTGATATGTTGGGTAGATTTTGGAGAAATCTGT 915
Qy 261 TyrSerLeuThrValProPheGlyGlnIleProAsnIleAspValThrAspAlaMetVal 280
Db 916 TACTCTTGGACAGTTCCCTTGGACAGAAACCAATGATGATGATGATGATGATGATGATGATG 975
Qy 281 AspGlnAlaTyrAspAlaGlnArgIlePheLeuGlyGlnIleGlyGlnIleGlyGlnIleGly 300
Db 976 GACCAAGCCCTGGAGATGCACAGAGATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1035
Qy 301 GlyLeuProAsnMetThrGlnIlePheTyrGluAsnSerMetLeuThrAspProGlyAsn 320
Db 1036 GGTCTTCTCAATATGATGATCAAGAGATTCGGGAAATTCATGCTAACGAGAGAGAGAGAG 1095
Qy 321 ValGlnIleValAlaCysHisProThrAlaTyrAspIleGlyGlyAspPheArgIle 340
Db 1096 GTTTCAGAAAG 1155
Qy 341 LeuMetCysThrIleValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 360
Db 1156 CTATATGTGCAAAAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
Qy 361 IleGlnTyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyValAsnGlu 380
Db 1216 ATCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275
Qy 381 GlyPheHisGlnAlaValGlyGlnIleMetSerLeuSerAlaIleThrProHisIleu 400
Db 1276 GGAATTCATGAG 1335
Qy 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnIle 420
Db 1336 AATTCATGTTGGTCTTCTGACCCGATTTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
Qy 421 LeuLeuGlyGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluGly 440
Db 1396 CTGCTCAAAAG 1455
Qy 441 ThrArgTyrMetValPheLeuGlyGlnIleProLysAspGlnIleThrPheLeuIleThrPhe 460
Db 1456 TGGAGGTGATGCTCTTTTAAAGGAGAAATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
Qy 461 GluMetLysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCys 480
Db 1516 GAGATGAG 1575
Qy 481 AspProAlaSerLeuPheHisValIleSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db 1576 GACCCGAGCTCTGTTTCAATGTTTAAATGATCTATCTATCTGATGATGATGATGATGATG 1635
Qy 501 ThrLeuTyrGlnPheGlnPheGlnGlnIleAlaLeuCysGlnAlaAlaIleHisIleGluPro 520
Db 1636 ACCCTTACCAATTCAGATTCAG 1695
Qy 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnIleLeuPheAsnMetLeu 540

```

RESULT 3
US-10-114-893-85
; Sequence 85, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racle, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 3325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-85

Alignment Scores:
Pred. No.: 0
Score: 3865.00
Percent Similarity: 100.00%

Length: 3325
Matches: 719
Conservative: 1

```

QY	1	SerThrIleGluIugInalAlvSthPhleuaspLyPheasnHLSglValaIuasp	20
Db	73	TCACACCTTAGGAAACAGGCCAGACATTTTGGACAGATTMACCGAACCGAAGAC	133
QY	21	LeuPheyrInIuSerSerLeuAlaSerTPrasntYrsnthRsnIlethrgIuIasn	40
Db	133	CTGTCTCATPAAAGTGCACCTTGCTCTTGGAAATTATTAACCAATATTACTGAAAGAAAT	192
QY	41	ValGlnAsnMetAsnAsnAlaGlyAspLyStrSerLaphleuysglInuSerThr	60
Db	193	GTCCAAAACATGAAATATGCTGGGGCAAAATGGCTCTCTTTTAAAGNACGTCACCA	252
QY	61	LeuAlaGlnMetLytrProleuGlnIuIleGlnAsnLeuthrValLysleuGlnLeuGln	80
Db	253	CTTGCCCAAAATGATCCACTACACAAAGAAATTCACAAATCTCACTCAAGCTTCAGCTGCAG	312
QY	81	AlaLeuGlnGlnAsnGlySerSerValIleSerGlnAspLySerLysAspGlyLeuAsnthr	100
Db	313	GCCTTCAGCAAAATGGCTTTCAGTCTCTCGAAAGACAGAACCAAGCTTGAACACA	372
QY	101	IleLeuAsnthrMetSerThrIleYrSerThrGlyLysValCysAsnProAspAsnPro	120
Db	373	ATTCTTAAATACAAATGAGCACCACTCTACAGTACTGGAAAGTTGTAAACCAATATTCACA	432
QY	121	GlnGlnCysLeuLeuLeuLeuGlnProGlyLeuAsnGlnIleMetValAsnSerLeuAspTytr	140
Db	433	CAAGATGCTTATTTACTTGACACAGGTTTGATGAATATATGGCAACAGTTTAGACTAC	492
QY	141	AsnGlnAspLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAspPro	160
Db	493	AATGAAAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGCTGGCAAGCAGCTGAGGCCA	552
QY	161	LeuTytrGlnGlnIleTytrValValLeuLysAsnGlnIleMetAlaAsnHISLytrGluAsp	180
Db	553	TATATGAAAGATATATGGCTCTTGAANAATGAGATGGCAAGCAAAATCATTAATGAGGAC	612
QY	181	TytrGlyAspTytrTrpArgGlyAspTytrGluValAsnGlyValAspGlyTytrAspTytrSer	200
Db	613	TATGGGATTAATTGGAGAGAGAACTATGAAGTAATAATGGGGTAGATGGCTATGACTACACG	672
QY	201	ArgGlyGlnLeuIleGlnAspValGlnHISThrPheGlnGlnIleLysProLeuTytrGlu	220
Db	673	CGCGGCCACCTTGATTTGAAGATGTGGAACTACCTTTGAAAGAAATTAACCATTAATATGAA	732
QY	221	HISLeuHISAlaTytrValArgAlaLysLeuMetAsnAlaTytrProSerTytrIleSerPro	240
Db	733	CATCTTCATGCCTATATGTAGAGGCAAAAGTTGATGAATGGCTATCTTCATATACGTCA	792
QY	241	IleGlyCysLeuProAlaHISLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeu	260
Db	793	ATTGGATGCTCCCTGCTCATTTGGTGTGTATGTGGGGTAGATTTTGGCAAAATGTG	852
QY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
Db	853	TACTCTTTGACAGTCCCTTGGACAGAAACCAACATAGATGTACAGATGCCAATGGTG	912
QY	281	AspGlnAlaTrpAspAlaGlnArgLlePheLysGlnAlaGluLysPhePheValSerVal	300
Db	913	GACCAAGGCTGGGATGACACAGAAATATTCAAGAGGCGGAGAAATCTTGTATCTGATT	972
QY	301	GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320
Db	973	GGTCTCTCTAATATGACTACAGAGATCTCGGAAAAATTCATGCTTAACGACCCAGGAAT	1032
QY	321	ValGlnLysAlaValCysHISProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle	340
Db	1033	GTTCAGAAAGAGTCTGGCCATCCACAGCTTGGGACCTGGGGAAGGCGGACTTCAGAGTTC	1092

OY	341	LeuMetCysThrIysValIleTrmMetAspPheLeuThrIleHisSHISGLIuMetGlyHis	360
Db	1093	CTTATGTGCACAAAGGTGCACAAATGGAGACTTCTGCAGCTCATCTGAGATGGGCAT	1152
OY	361	IleGlnIlyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGlu	380
Db	1153	ATCCAGATGATGATGGCATATGGTGCACAACCTTTCTGCTAAGAAATGGAGCTAATGAA	1212
OY	381	GlyPheHisGluAlaValIleGlyGluIleMetSerLeuSerAlaIleThrProIlyHisLeu	400
Db	1213	GGATTCATAGAAAGTGTGGGAAATATGCACTTTCTCGAGCCACACCTTAAGCAATTTA	1272
OY	401	LysSerIleGlyLeuLeuSerProAspPheGlnIluAspAsnGluTrpGluIleAsnPhe	420
Db	1273	AAATCCATTTGGTCTTCTGTCACCCGATTTTCAAAAGACATGAAGAGAAATAACTTC	1332
OY	421	LeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGluLys	440
Db	1333	CTGCTCAACACAGACACACAGCATTTGGGCTCTGCACTTACTTTCATGATTGAGAAAG	1392
OY	441	TyrTrpArgPheMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrp	460
Db	1393	TGGAGGTGGATGTCCTTTAAAGGGGAAATTTCCAAAGACAGTGGATGAAAAGTGGTGG	1452
OY	461	GluMetLysArgGluIleValIleGlyValIleGluProValProHisAspGluThrTyrCys	480
Db	1453	GAGATGAAGCAGAGATAGTGTGGGTGGTGAACCTGCGCCCATGATGAACAATACGT	1512
OY	481	AspProAlaSerLeuPheHisValIserAsnAspTyrSerPheIleArgTyrTyrThrArg	500
Db	1513	GACCCCGCATCTGTTCCATGTTTCAATGATTACTCATTCATTCATTCATTTACACAAAG	1572
OY	501	ThrLeuTyrGlnPheGlnPheGlnAlaLeuLysGlnAlaIleLysHisGluGlyPro	520
Db	1573	ACCTTTACCAATTCAGATTTCAGAAAGCACTTTGTCAAGAGAGTAAACATGAAGGCCT	1632
OY	521	LeuHisLysCysAspIleSerAsnSerThrGluAglGlyIleuPheAsnMetLeu	540
Db	1633	CTGCACAATGTGACATCTCAACCTCAACAACACTGGACAGAAACGTTCCAAATATCTG	1692
OY	541	ArgLeuGlyLysSerGluProTyrThrLeuAlaLeuGluuAsnValIleAlaLysAsn	560
Db	1693	AGCGTTGAAATACAGAACCTTGACCTTACCATTTGGAAATGTGTTGAGGACAAAGAAC	1752
OY	561	MetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGln	580
Db	1753	ATGATATGTAGGCACCTGCTCAACTACTTTGAGCCCTATATTACCTGGCTGAAGACAG	1812
OY	581	AsnLysAsnSerPheValIleGlyTyrPserThrAspTrpSerProTyrAlaAspGlnSerIle	600
Db	1813	AACAGAAGATTTTGTGGGATGAGTACCCACATGGAGTCCATATGCGACACCAAGATC	1872
OY	601	LysValAlaGlnIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsn	620
Db	1873	AAAGTGAGNATTAAGCTTAATAATCAGCTTTGGAGATTAAGCATATGATGAGACAGACAT	1932
OY	621	GluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVal	640
Db	1933	GAAATGTACCTGTTCCGATCATCTGTTCCATATGCTATAGAGCAGGTACTTTTAAAGTA	1992
OY	641	LysAsnGlnMetIleLeuPheGlyValuGluAspValArgValAlaAsnLeuLysProArg	660
Db	1993	AAAAATCAGATGATCTTTTGGGAGAGAGATGTGCCAGTGGCTAAATTTGAACCAAGA	2052
OY	661	IleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThr	680
Db	2053	ATCTCCCTTAATTTCTTGTCTACGTGCACCTTAATAATGTCTGATATCATCTCTAGAACT	2112
OY	681	GluValIleGlyLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsn	700
Db	2113	GAAGTGAAGAGCCATGAGATGTGCCGGAACCGTATCATAGATGCTTCCGTGTGAT	2172

```

OY 701 AspAsnSerLeuGluPheIleuGlyIleGlnProThrLeuGlyProPheAsnGlnPro 720
DB 2173 GACAAAGCCTAGAGTTCTGGGGATACAGCAACCTTGGACCTCTTAACGAGCCCT 223

RESULT 4
US-09-969-384-2
: Sequence 2, Application US/09969384
: Publication No. US20020192749A1
: GENERAL INFORMATION:
: APPLICANT: Moore, et al.
: TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
: FILE REFERENCE: PT055P1
: CURRENT APPLICATION NUMBER: US/09/969,384
: PRIOR FILING DATE: 2001-10-03
: PRIOR APPLICATION NUMBER: PCT/US01/10542
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 60/236,384
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/194,118
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 2920
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1707)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2702)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2749)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2757)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2788)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2789)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2819)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2835)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (2856)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-2

Alignment Scores:
Pred. No.: 0 Length: 2920
Score: 3833.00 Matches: 715
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.17% Mismatches: 4
Query Match: 99.07% Indels: 1
DB: 9 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-09-969-384-2 (1-2920)
OY 1 SerThrIleGluGlnAlaIysThrPheLeuAspIysPheAsnHisGluAlaGluAs 20
DB 83 TCCACCACTTGAAGAAAGGCGCAAGACATTTTGGGAGCAAGTTTAACCGAAGCCGAGA 142
OY 20 PleuPheTyrGlnSerSerIleuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu 40
DB 143 CCGTTCTATCAAGATTCACCTGCTCTTGGGAATTATTAACCAACATATTACTGAAGAGA 202

```

QY 40 nValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerTh 60
 Db 203 TGTCCAAAACATGATTAATGCTGGGACAAATGCTGCTCTTTTAAAGAAAGCTCAC 262
 QY 60 rleuAlaGlnMetLysTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
 Db 263 ACTGTCCCAAAATGATTCACATCAAGAAATTCAGAAATCTCACAGTCAAGCTTCACTGCA 322
 QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
 Db 323 GGCCTTCAGCAAAATGGGTCTTCAGTCTCTCAGAAACAGCAAGCAACCGTTGACAC 382
 QY 100 rIleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPr 120
 Db 383 AATTCATAATACAAAGACACCATCTACAGTACTGAAAGATTGTAAACCCAGATATACC 442
 QY 120 oGlnGlnCysLeuLeuLeuGlnLupProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
 Db 443 ACAAGATGCTTATTAATGTAACCGAGTTTGAATGAATGAATGCAAAACAGCTTAACTA 502
 QY 140 rAsnGlnArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPr 160
 Db 503 CAATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTCCGCAAGCAGCTGAGGCC 562
 QY 160 oLeuTyrGlnGluTyrValIvalLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAs 180
 Db 563 ATTATATAGAGATGATGTGCTCTGAAAAATGAGATGGCAAGACCAATCATTTATGAGA 622
 QY 180 PTYrGlnAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSe 200
 Db 623 CTATGGGATTAATTTGAGAGAGAGACTATGAATGAATGGGTAGAGTGTATGATGACG 682
 QY 200 rArgGlnGlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrG 220
 Db 683 CCGCGGCGAGTGTGATGAAGATGTGGAACATACCTTTGAAGAGATTAACCATTAATGA 742
 QY 220 uHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPr 240
 Db 743 ACATCTTATGCTATGTGAGGCCCAAGATGATGAATGCTATCTCTCATTAACATGCTC 802
 QY 240 oIleGlyCysLeuProAlaHisLeuLeuGlnLysAspMetTrpLysArgPheTrpThrAsnLe 260
 Db 803 AATTTGATGCTCCCTGCTCATTTGCTTGGATATGATGGGCTGATTTGGACAAATTT 862
 QY 260 uTyrSerLeuThrValProPheGlnGlnLysProAsnIleAspValThrAspAlaMetVa 280
 Db 863 GTACMSTTTGACAGTTCCTTCTTGACAGAAACCAACATAGATGTCTGATGCAATGGT 922
 QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerVa 300
 Db 923 GGACCAAGRCTGGGATGACAGAGAAATATTCAGAGAGCGGAGAAAGTTCTTTGTATCTGT 982
 QY 300 lGlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAs 320
 Db 983 TGGCTCTCCATTAATGACTCAAGATCTGGGAAATTCATGCTAACGACACCGAGAAA 1042
 QY 320 nValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgI 340
 Db 1043 TGTTCAGAAACCACTGCGCATCCACAGCTTGGGACCTGGGAGGCGCACTTCAGAT 1102
 QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGlnGlnMetGlyH 360
 Db 1103 CCTTATGTGCAGAAAGGTGACATGACGACTTCTCAGACCTCATATATAGATGGGCA 1162
 QY 360 sIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnG 380
 Db 1163 TATCCAGATGATATGCAATATGCTGCACAACCTTTCTCTAAGAAATGGAGCTATATGA 1222
 QY 380 uGlyPheHisGlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLe 400
 Db 1223 AGATTCATGATGAACCTGTGGGAAATCATGCTACTTTCGACGCGCACACCTTAAGCATTT 1282

QY 400 uLysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPh 420
 Db 1283 AAAATTCATTGGTCTTGTGACCCCATTTTCAAGAGACAAATGAACAGAAATTAACCTT 1342
 QY 420 eLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGly 440
 Db 1343 CTGTCTCAAAACAGCACTCAGATGTGTGGAGCTGTGCCATTTACTTACATGTTACAGAA 1402
 QY 440 sTrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpTr 460
 Db 1403 GTGAGGTGGATGGCTTTTAAAGGGGAAATTCCTCAAAAGCACTGATGATAAAAGTGTG 1462
 QY 460 pGluMetLysArgGlnIleValGlyValValGluProValProHisAspGluThrTyrCy 480
 Db 1463 GGAGATGAACGAGAGATATGTTGGGTGGTGGAGACCTGTGCCCATGATGAACATACTG 1522
 QY 480 sAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrAr 500
 Db 1523 TGACCCCGCACTCTCTGCTCCATGTTCTAATGATTAATCATTCATTCATATTTACACAG 1582
 QY 500 gTrpLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlyLys 520
 Db 1583 GACCCCTTACCAATTCAGTTTCAGAGAGCACTTTGTCAAGCAGCTTAAACATGAAGGCC 1642
 QY 520 oLeuHisLysCysAspIleSerAsnSerThrGlnAlaGlnGlyLysLeuPheAsnMetLe 540
 Db 1643 TCTGCAGAAATGAGATCTCAAACTCTACAGAGCTGGACAGAACTTTCATATGTCT 1702
 QY 540 uArgLeuGlyLysSerGluProTrpPheLeuAlaLeuGlnLysAsnValGlyAlaLysAs 560
 Db 1703 GAGGNTTGGAAATTCAGAACCCCTGACCTGACATTTGGAATGTTGTGAGCAAGAA 1762
 QY 560 nMetAsnValArgIleSerLeuLysSerLysPheGlnPheGlnPheTrpLeuLysAspG 580
 Db 1763 CATGATGTGAAGCCACTGCTCAACTTGTAGCCCTTATTTACTGCTGTAAGACCA 1822
 QY 580 nAsnLysAsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerI 600
 Db 1823 GAACAAGAAATCTTTTGTGGAGTGAAGTACCGACCTGAGTGCATATGACCAAGCAT 1882
 QY 600 eLysValArgIleSerLeuLysSerLysAlaLeuGlnLysAspLysAlaTyrGluTrpAsnAsp 620
 Db 1883 CAAGTGAAGATTAAGCTTAATCACTCTTGGAGTAAAGATTAAGATGAATGAACAGCAA 1942
 QY 620 nGluMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysVa 640
 Db 1943 TGAATGTACCTGCTCCATCATCTGTTCCATATGCTATGAGACGACTTTTAAAAGT 2002
 QY 640 lLysAsnGlnMetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProAr 660
 Db 2003 AAAAATTCAGATGATCTTTTGGGGAGAGATGTGGAGTGTGAAATTTGAAACCAAG 2062
 QY 660 gIleSerPheAsnPhePheValThrAlaProLysAsnValSerAspIleLeuProArgTh 680
 Db 2063 AATCTCTTAAATTTCTTGTCTACCTGCACTTAAATGTGTGTATCATCTTCAGAAC 2122
 QY 680 rGluValGluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAs 700
 Db 2123 TGAAGTTGAAGAGCATCAGATGTCCGGAGCGGTATCATATGTTCCGTCTGAA 2182
 QY 700 nAspAsnSerLeuGlnPheLeuGlyLysIleGlnProThrLeuGlnLysProProAsnGlnProPr 720
 Db 2183 TGAACAGACCTTAGATTTCTGGGATACAGCAACACTTGGACCTCTTAACAGACGCC 2242
 QY 720 o 720
 Db 2243 T 2243

RESULT 5
 US-10-158-847-137
 ; Sequence 137, Application US/10158847
 ; Publication No. US20030091557A1
 ; GENERAL INFORMATION:

APPLICANT: Tom Parry et al.
TITLE OF INVENTION: Method and Compositions for Modulating ACE-2 Activity
FILE REFERENCE: PF557
CURRENT APPLICATION NUMBER: US/10/158,847
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: 60/295,004
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 158
SOFTWARE: PatentIn version 3.1
SEQ ID NO 137
LENGTH: 2920
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1707)..(1707)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2702)..(2702)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2749)..(2749)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2757)..(2757)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2788)..(2788)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2819)..(2819)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2835)..(2835)
OTHER INFORMATION: n equals any amino acid
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2856)..(2856)
OTHER INFORMATION: n equals any amino acid
US-10-158-847-137
Alignment Scores:
Pred. No.: 0
Score: 3833.00
Percent Similarity: 99.31%
Best Local Similarity: 99.17%
Query Match: 99.07%
Length: 2920
Matches: 715
Conservative: 1
Mismatches: 4
Indels: 1
Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x US-10-158-847-137 (1-2920)
QY 1 SerThrIleGluGluGlnAlaIysThrPheLeu-AspLysPheAsnHisGlnAlaGluAs 20
Db 83 TCACACATTGAGGAAAGCCAGCCAGACATTTTGGGACAAAGTTAAACCCAGAACCCGAGAA 142
QY 20 pLauPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAs 40
Db 143 CCGTCTTATCAAAAGTTCACCTTCCTTGGAAATTTAAACACCAAAATTAATGAAAGAA 202
QY 40 nValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuGluGlnSerThr 60
Db 203 TGTCCAAACATGAATATATCTGCGGACAAATGCTGCTTTTAAAGAAACAGTCCAC 262
QY 60 rLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
Db 263 ACTTGCCCAATGTATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTCAAGCTGCA 322

QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
Db 323 GGCCTTACGAAATAATGGGCTCTCAGTCTCCAGAAACAGAGCAACAGGTTGAACAC 382
QY 100 rIleLeuAsnThrMetSerThrIleTyrSerThrGlyValAlaLysAsnProAspAsnPr 120
Db 383 AATTCTAATACATGACACCATCTACAGTACTGGAAAAATTGTATACCCAGATATACC 442
QY 120 oGlnIleGlySerLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
Db 443 ACAAGAAATGCTTATCTTGAACACAGTTTGAATGAATATATGCAAAACAGTTTGACTA 502
QY 140 rAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlySerGlnLeuAlaGPr 160
Db 503 CAATGAGAGCTCTGGGCTGGGAAAGCTGAGATGAGTAAATGGGTAGATGCTATGACTACAG 562
QY 160 oLeuTyrGluGluTyrValValIleLeuLysAsnGluMetAlaArgAlaAsnHisLysTyrGluAs 180
Db 563 ATTATATGAAGATGTGTGCTTGAATAAATGAGATGGCAAGACCAATCTTATGAGGA 622
QY 180 pTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSe 200
Db 623 CTATGGGATATATGAGAGAGAGACTATGAAATGAGGTAGATGCTATGACTACAG 682
QY 200 rArgGlyGlnLeuIleGlyAspValGluHisThrPheGluGluIleLysProLeuTyrG 220
Db 683 CCGCGCCAGTGTATGGAAGATGGGAAACATTCCTTGAAGAGATTAACCATATATAGA 742
QY 220 uHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPr 240
Db 743 ACATCTTCATGCTTATGAGGCCAAAGTTGATGAATCCCTTCTCTTATACAGTCC 802
QY 240 oIleGlyCysLeuProAlaHisLysLeuGlyAspMetTrpGlyArgPheTrpHisnLe 260
Db 803 AATTGATGCCCTCCCTGCTCATTTGCTGGATATGCGGGATGATTTGGACAAATYT 862
QY 260 uTyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVa 280
Db 863 GTACWSTTTGACAGTTCCTTTGGACAGAAACCAACATGATGTATGATGCAATGCT 922
QY 280 lAspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheAlaSerVa 300
Db 923 GGACCAAGCTTGGATGACAGAGAAATTTCAAGAGGCCAGAAAGTTCTTGTATCTGT 982
QY 300 lGlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAs 320
Db 983 TGGTCTTCCTAATATGACTCAAGAGATCTGGGAAATTCATGCTAACGAGCCAGAGAA 1042
QY 320 nValGlnLysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheAlaG 340
Db 1043 TGTTCAGAAAGCATGCTCCCATCCACAGCTTGGGACCTGGGAGGCGACTTCAGAGAT 1102
QY 340 eLeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyH 360
Db 1103 CCTTATGTGCACAAAGGAGACATGAGAGACTCTCTGACAGCTCATCTAGATGGGGCA 1162
QY 360 sIleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnG 380
Db 1163 TATCCAGATGATGATGAGATATGCTGCACAACTTTTCTGTAAGAAATGAGGATATGA 1222
QY 380 uGlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLe 400
Db 1223 AGGATTCCTGAAGCTGTGGGGAATATGATGATCTTCTCAGCCACACTTAAGCAATTT 1282
QY 400 uLysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPh 420
Db 1283 AAAATCCATGTGCTTCTGCTACCCGATTTTCAAAAGACATGAAAGAAATTAATATT 1342
QY 420 eLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGly 440
Db 1343 CCTGCTCAAAACAGACACTCAGCATTTGTTGGACTCTGCCATTTACTTACATGTTAGAGAA 1402
QY 440 sTrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTr 460

```

Db 1403 GTGGAGGTGGATGCTTTAAAGGGGAATTCACCAAGACGATGAGAAAGTGGTG 1462
QY 460 pglumetlysargluilevalglvalvalgluprovalprohisaspgluthtyrcy 480
Db 1463 GGAGATGAAGCAGAGATAGTGGGGTGTGGAACCTGTGCCCATGTATGAACATATCG 1522
QY 480 saspaprolasertleuphehisvalserasnasphtyrspherhellatgtyrtyrthar 500
Db 1523 TGACCCCGCATCTGTTCCATGTTTCAATGATTACATTCATTCATTCATTCATTCACAAAG 1582
QY 500 gthleutyrglnpheglnpheglnlglualaleucysglnlaalaalyslsglulypyr 520
Db 1583 GACCTTTCACCAATTCAGATTCCAGAAAGCACTTTGTACAGCAGCTTAACATGAGGCC 1642
QY 520 oleuhsilyscysaspiliserasnserthrghuallglulglulysleupheasnmetle 540
Db 1643 TCTGCACAAATGTGACATCTCAACTACAGAAAGCTGCAGAAACGTTCCAAATATGCT 1702
QY 540 uargleuglylyserserluprotptthleualaleugluasnvalvalglvalalysas 560
Db 1703 GAGGWTGGAAATCAGAACCCCTGCACCTAGCATGGAAATGTTGTAGAGCAAGAA 1762
QY 560 nmetasnvalargproleuasntrpgheluproleuphehtrpgleuysaspgl 580
Db 1763 CATGAATGTAGGCGCATGCTCACTACTTGTAGGCCCTTATTTACCTGGCTGAAAGACCA 1822
QY 580 nasnlysasnserphevalgltyrtpserthrastpserprotyralaaspglinser1 600
Db 1823 GAAACAAGATTTCTTTGTGGGATGAGATGACGACGAGTCGCATATGACACCAAGCAT 1882
QY 600 elysvalargliserleuylserseraleuglyaspysalatyrglutrpasnaspas 620
Db 1883 CAATGTGAGGATTAACCTTAATCAGCTCTTGAGATTAACCATATGATGAACGACCA 1942
QY 620 nglumetlyrleupheargseraserValalatyralametrarglntrpheleuysa 640
Db 1943 TGAATGTACTCTGTTCCATCATCTGTGCAATAGCTATGAGGAGTACTTTTAAAGT 2002
QY 640 llyasnglumetileleuphegllygluluaspyalargvalalasnleuysproar 660
Db 2003 AAAAATCAGATGATTTCTTTGGGAGAGAGATGGAGTGGCTAATTTGAACCAAG 2062
QY 660 gliserpheasnpherphevalthrallaproluasnvalseraspilleproagth 680
Db 2063 AATCTCCTTAATTTCTTTGTCACCTGCACCTTAAATGTGCTGATATCATTTCTTAAAC 2122
QY 680 rgluvalglulysalaleargmetseraserArgllesnasaplapheargleuas 700
Db 2123 TGAAGTTAAAAAGCCATCAGAGATGTCGCCGAGCCGATTCATATGCTTTCCGCTGAA 2182
QY 700 naspasnserleuglnpheuglylleglnprothleuglyproproasnnglnpropr 720
Db 2183 TGACAGACAGCTAGAGTTCTTGCGGATACAGACCAACCTTGACCTCTTAACGAGCCCC 2242
QY 720 o 720
Db 2243 T 2243

```

```

RESULT 6
US-09-969-384-12
; Sequence 12, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: P05551
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCY/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/2236,384
; PRIOR FILING DATE: 2000-09-25

```

```

; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1705)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-12

Alignment Scores:
Pred. No.: 0 Length: 2911
Score: 3785.00 Matches: 715
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.17% Mismatches: 4
Query Match: 97.83% Indels: 3
DB: 9 gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x US-09-969-384-12 (1-2911)
QY 1 SerThrIleGluGlnAlaLysThrPheLeu-AspLysPheAsnHisGluAlaGlns 20
Db 83 TCACACATTAGAGACAGCCGCAACATTTTGGACAAAGTTTAACCAAGCCGAAAGCA 142
QY 20 PleuPheTyrglnSerSerleuAlaSerTrpAsnTrpAsnIleThrGluGlns 40
Db 143 CCGTCTCATCAAAAGTTCACTGCTCTTGGAATTAACCAATATTAACGAGAGA 202
QY 40 nValGlnAsnMetCasnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnSerTh 60
Db 203 TGTCCAAACATAAATTAATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCAC 262
QY 60 rLeuAlaGlnMetTyrrProleuGlnIleGlnAsnLeuThrValLysLeuGlnGlu 80
Db 263 ACTTGCCCAATATATCCACTACACAAATTCGAATTCACAGTCACACTTCAGTGA 322
QY 80 nAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnTh 100
Db 323 GGCTCTTACCAAAATAGGGCTTCAGTGTCTTCAGAAAGCAAGAGCAACGGTTGAAC 382
QY 100 rIleLeuAsnThrMetSerThrIleTyrrSerThrGlyLysValCysAsnProAspAsnPr 120
Db 383 AATTCCTAAATACATAGACACCACTACAGTCTGGAAGAAAGTTGTAAACCCAGATATCC 442
QY 120 oGlnGluCysLeuLeuLeuGlnIupProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTy 140
Db 443 ACAAGATGCTTATTTACTTGAAACCAAGTTGAATGAATTAATGCGCAACAGTTTAGCTA 502
QY 140 rAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPr 160
Db 503 CAATGAGAGGCTCTGGGCTTGGGAAGCTGAGATCTGAGTGGCGGAACACACTGAGGCC 562
QY 160 oleuTyrgluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrgLus 180
Db 563 ATTATATGAAGATAGTGTCTTGAATAATGAGTGGCAAGCAACAAATCATATAGAGA 622
QY 180 pTyrgLysPtyrTrpArgGlyAspTyrgLualValAsnGlyValAspGlyTyrgLysSe 200
Db 623 CTATGGGATTTATTGAGAGCGAGACTATGAAAGTAATGGGATGAGTGGCTATGACTTAAG 682
QY 200 rArgGlyGlnLeuIleGluAspValGluIsthrPheGluGlnIleLysProLeuTyrgl 220
Db 683 CCGCGGCGAGTTGATTAAGATGTGAACATACCTTTGAAGAGATTAAACCTTATATGA 742
QY 220 uHisLeuHisAlaTyrrValAlaArgAlaLysLeuMetAsnAlaTyrrProSerTyrrIleSerPr 240

```


PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090445
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090538
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:
Score:	3720.00	717	1	2	138
Best Similarity:	83.88%				
Best Local Similarity:	83.76%				
Query Match:	96.15%				
Gaps:					1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-028-072-71 (1-3732)

QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCACCATTTGAGAGACAGCCAGACATTTTGGACAGATTACCCAGAGCCGAGAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluGluAsn 40
 DB 154 CTGTCTATCAAGAGTCACTTGTCTTGTGAATTAACACCAATTTCTGAAGAGAT 213
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAAGATGATATCTGCGGACAAATGCTGCTTTTAAAGAGACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCAAATGATCATCTACAGAAATTCAGAAATTCACAGTCAAGCTTCAAGCTGCGAG 333
 QY 81 AlaLeuGlnGlnAsnLysSerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTCTTCAGCAAAATGGGCTTCAGTCTCAGTCAGAGACAGCAAGCAAGGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnPro 120
 DB 394 ATTCATAATACATGAGACACATCTACAGTCTGAGAAAGTTGTAACCCAGATATACCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAAGATGCTTATCTTACCTGACCAAGTTGAATGAATATATGCAACAGTTTACACTAC 513
 QY 141 AsnGluArgLeuTyrAlaTyrPheLysSerTyrPargSerGluValGlyLysGlnLeuArgPro 160
 DB 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTGGCAAGAGAGTGAAGCCA 573
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TATATGAAGAGATGTGCTTGAATAATGAGATGCAAGAGCAAAATCATATATGAGGAC 633
 QY 181 TyrGluAspTyrTyrPargLysAspTyrGluValAsnGluValAspGlyTyrAspTyrSer 200
 DB 634 TATGGGATTTTGGAGAGAGACATCTGAAAGTAAATGGGTGAGATGCTATGACTACAC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CGGGGCGAGTGTATGAGATGAGAAACATACCTTGAAGAGATTAAACCATTAATATGAA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTTCATGCTATGTGAGGCAAAAGTTGATGAAGCTTATCTCATATCACTGATCCA 813
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrAsnLeu 260

DB 814 ATTGATCCCTCCCTGCTCATTTGCTTGCTGATATGCGGGTACATTTTGGACAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACCTTTGAGAGTTCCTCTTGGACAGAAACCAATAGATGTACTGATGCAATGTG 933
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACAGGCGCTGGAGTACAGAGATATTCAGAGAGCGCCGAGAGTCTTGTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrGluAsnSerMetLeuTyrAspProGlyLys 320
 DB 994 GGTCTCTTATATGATGACCAAGATTTCTGGAAATTCATCTATACGAGACCGAGAAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAAGAGTCTGTCATCCACAGCTTGGGACCTGGGAGAGGGGAGCTTCAAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHis 360
 DB 1114 CTATAGTCCAGAAAGTGCACATGACACATCTCTGACAGCTCATCATGAGATGGGCAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATATGCTATATGCTGACACACCTTTCTGCTTAAGAAATGGAGTAAATGA 1233
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GGATTCATGTAAGCGTGGGGAATATCATGTCACCTTCTGAGCCACACACCTTAACATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1294 AAATCCATTTGGCTCTCTGCTCAACCGATTTTCAAGAGACATGAAAGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1354 CTGCTCAACAGACATCATGATTTGGGCTCTGCGCATTTACTTTCATGATTTAGAGAAAG 1413
 QY 441 TyrPargTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysLysTyrP 460
 DB 1414 TGGAGGTGAGAGTCTTTTAAAGGGGAATTCACAAAGACGCTGAGTGAAGAAATGGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCys 480
 DB 1474 GAGATGAAGCGAGAGATGATGGGTGGTGAACCTGCGCCCATGATGAACATACTGCT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnSerTyrSerPheIleArgTyrTyrArg 500
 DB 1534 GACCCGAGTCTCTGCTCATGTTCTGATGATATTCATTCATTCATTCATTCATTCATTC 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyPro 520
 DB 1594 ACCCTTTCACATTCACATTTCAAGAGACACTTTGTCAAGAGAGCTAAACATGAAGGGCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAACCTCTACAGAACTGACAGAAAGCTGT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCTCAAAATGTGAACCTCTCTAGATTTCAATTTACATTTCCATTCGCTTAGTTGG 1772
 QY 537 ----- 537
 DB 1773 TATTGATTTCTTTCTCTAAAGAAATTTTATGGCTCAAAATGCTCATTTTACAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTAATTTGTGTGACAGAGAACTTACACATCAACAAATTTGGGTGGGCCA 1892
 QY 537 ----- 537
 DB 1893 CCTCTTTCTCCCTATCAATTAACAGCCCTCTCTCTGATTTGGAAGAAAGAGCG 1952

QY 537 ----- 537
 Db 1953 GTTAAAGGTGGAATATCTGTTAAATGACATCTTTCTTATCTGCCAGAACAAATTT 2012
 QY 537 ----- 537
 Db 2013 AGCCAAAGTCAAGAGAGAAACCATGATCATGATGTAATATATGACATCTGAGAAC 2072
 QY 538 ----- AsnMetLeuArgLeuGlyLysSer 545
 Db 2073 CCTCAAAAGGCCCTGACCCCTTTTGTGTGACAAATATCTGAGGCTTGGAATAATCA 2132
 QY 546 GIUPTPTPTThrLeuAlaLeuGluAsnValValAlaGlyAsnMetAsnValArgPro 565
 Db 2133 GAACCTGGACCCCTGACATGGAATAATGTTAGAGAC- AAGAACATGAATGTAAGGCUCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 Db 2192 CTGCTCACTACTTTGAGGCCCTTATTACCTGGCGAAGACAGAACAGAAATCTTTT 2251
 QY 586 ValGlyTyrSerThrAspTrpSerProTyrAlaAsp- GlnSerIleLysValArgIleSe 605
 Db 2252 GTGGGATGAGTACCGACTGAGTCCATATGACAGACCAAGCAAGTAGAGATTAAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
 Db 2312 CCTAAATCAGCTCTTGAGATAAAGCATATGAGAACAGCAATGAATGTACCTGTT 2371
 QY 625 eArGSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
 Db 2372 CCGATCATCTGTTGATATGATATGAGCAGTACTTTTAAAGTAAATAATCATGAT 2431
 QY 645 eleuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 Db 2432 TCTTTTGGGGAGGAGATGTGCGAGTGGCTTAATTGAAACCAAGAAATCTCTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluVal 685
 Db 2492 CTTTTCACCTGACCTTAAATAATGTCTGATATCTTCTGAACTGAAAGTTGAAAGGC 2551
 QY 685 aIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 Db 2552 CATCAGGATGTCGCCGAGCCGATATCATGATGCTTCGCTGATATGACACACGCTTGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2612 GTTCTGGGAGATACAGCCACACTTGACCTCTTAACGACGCCCT 2657
 RESULT 8
 US-10-121-049-71
 ; Sequence 71, Application US/10121049
 ; Publication No. US2003002239A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C17
 ; CURRENT APPLICATION NUMBER: US/10/121,049
 ; CURRENT FILING DATE: 2002-04-12
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-121-049-71
 Alignment Scores:
 Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1
 US-09-978-385-2_COPY_19_738 (1-720) x US-10-121-049-71 (1-3732)
 QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 Db 94 TCACACATGAGAGAGACAGCCAGACATTTTGGACAAGTTTAAACCAAGCCGAGAAC 153
 QY 21 LeuPheTyrGlnSerSerIleuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
 Db 154 CTGTTCTATCAAAAGTTCACCTGCTGCTGGAATTAACCAACCAATATCTGAAAGAGAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
 Db 214 GTCCAAACATGATATATGCTGGGACAAATGTCTGCTTTTAAAGAACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 Db 274 CTGGCCCAATGATATCCATCAAGAAATTCAGAAATCTACACATCAAGCTTCAAGCTGACG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 Db 334 GCCTTCACGAAATGGCTGCTTCACAGCTCTCAGAACAGAACAGCAAGCGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAsnPro 120
 Db 394 ATTCTAATATCAATGACACCATCTACAGTCTGGAAGTGTGTAACCAATATCA 453
 QY 121 GlnGlnCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 Db 454 CAAGAAATGCTTATATCTGTAACCAAGTTTGAATGAATATGCAACAAAGCTTAACTAC 513
 QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
 Db 514 AATGAGAGGCTCTGGGCTTGGGAAACCTGAGATCTGAGCTGGCCAAAGCAGCTGAGGCCA 573
 QY 161 LeuTyrGluGluTyrValValIleLysAsnGluMetAlaArgAlaAsnHisLysTrpGluAsp 180
 Db 574 TTATATGAAGAGATATGTGCTTGAATAATGAAATGCAAGCAATATATATATGAGGAC 633
 QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyLysAspTyrSer 200
 Db 634 TATGGGATATATGAGAGAGACATGAAAGTAAAGGGGTATGATGCTATATGACTACAGC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
 Db 694 CGCGGCCAGTGTATGAAGTGTGAACATACCTTTGAAGAATTAACCATTAATATGAA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysIleMetAlaAlaTyrProSerTyrTrpIleSerPro 240
 Db 754 CATCTTCAAGCTATATGAGGAGCAAAATGATGAATGCTATCTCTATATACGTCCA 813
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrPheAsnLeu 260
 Db 814 ATTGATGCTCTCCCTGCTCATTTGCTTGGTATATGTGGGTATGATTTTGGCAAAATCTG 873

QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 Db 874 TACTCTTGACAGTTCCTTGGACAAACCAACATAGATGTACTGTAGTGAATGGTG 933
 QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerVal 300
 Db 934 GACCAAGCCCTGGATGACAGAGAAATTTCAAGAGAGCCAGAAAGTTCTTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 330
 Db 994 GGTCTCTATATAGACTCAAGATCTGGGAAATTCATGCTAACGAGCCAGGAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheValIle 340
 Db 1054 GTTCAGAAAGCAGCTCCATCCACAGCTGGGAAAGCCGCTTCAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlyMetGlyHis 360
 Db 1114 CTATATGTGCACAAAGGACAGATGACAGACTTCTGACAGCTCATCTAGATGGGCAAT 1173
 QY 361 IleGlnTrpAspMetAlaTrpAlaAlaGlnProPheLeuAsnArgAsnGlyAlaAsnGlu 380
 Db 1174 ATCCAGATGATATGCGATATGCTGCACAACTTCTGCTAAGAAATGAGGTAATGAA 1233
 QY 381 GlyPheHisGlnAlaValIleGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 Db 1234 GGATTCATGAAAGCTGTGGGAAATCATCTCTCAGCCACCACTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnLysAsnGlnuThrGlnIleAsnPhe 420
 Db 1294 AATTCATGCTGTCTGTGTCACCGGATTTTCAGAAACAAATGAAACAGAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrIleMetLeuGlyLys 440
 Db 1354 CTGCTCAAAACAGACACTCAGATGTGGGACCTGCGCATTTACTTACATGTTAGAGAAC 1413
 QY 441 TrpArgTrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpTyr 460
 Db 1414 TGGAGGTGAGTGGCTTTAAAGGGGAAATTCACAAACCAAGTGAAGAAAGTGTGG 1473
 QY 461 GlnMetLysArgGlnIleValIleValIleGlnProValProHisAspGlnThrTyrCys 480
 Db 1474 GAGATGAAGCAGAGATAGTGGGGTGTGGAACCTGTGCCCATGTGAACATACTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 Db 1534 GACCCCGCATCTGTTCATGTTCTGATGATATCATTCATTCATATTTACACAAAG 1593
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGlnAlaLeuGlyGlnAlaLysHisGlyLysPro 520
 Db 1594 ACCCTTAAACCAATTCAGTTCAAGAAAGCACTTTGTCAAGCACTTAACATGAAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPhe 537
 Db 1654 CTGCACAAATGTGACATCTCAACTCTACAGAACTGTGACAGAACTGTT-GRAGAAT 1712
 QY 537 537
 Db 1713 ACCTCAAAATGTGACCTCTCTAGATTCAGATTAATCATTTCCATGCTAGGTTTG 1772
 QY 537 537
 Db 1773 TATTTGATTTCTTGTCTTAAAGAAATTTATGCTCAAAATGTCTCATTTACAA 1832
 QY 537 537
 Db 1833 ACCAAACATTTAATTTGTGTCAGACAGAACTAGACATACAAATTTGGTGGGCCA 1892
 QY 537 537
 Db 1893 CCTCTTTCTCCATATACATACTAGAGCCCTCTCTCTGTAATGGAAGAAAGAGCG 1952
 QY 537 537

Db 1953 GTTTAGGTGGAATATATCTGTTAATGATGATCTTTCTTATCTGCCAGAAACAAATTT 2012
 QY 537 537
 Db 2013 AGCCAAAGTCAAGAGAGAAACCATATGATCATATGATATATATATGATCATCTGGAAC 2072
 QY 538 538
 Db 2073 CCTCAAAAGCCCTGAAACCCCTTTTGTGTAGCAATATCTGAGGCTTGAAATCA 2132
 QY 546 GlnProTrpThrLeuAlaLeuGlnValIleGlyAlaLysAsnMetAsnValArgPro 565
 Db 2133 GAACCTGGAACCTTACATTTGAAATGTTGAGAGC- AAGAACATGAAATTAAGGCCA 2191
 QY 566 LeuLeuAsnTrpPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 Db 2192 CTGCTCACTACTTTAGCCCTTATTTACTGTGCTGAAGACCAAGACAGAAATTTCTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleAsp 605
 Db 2252 GTGGATGAGAGTACGAGTGCATATGCAAGCCAAAGCATCAAGTGAAGATAG 2311
 QY 605 TrpLysSerAlaLeuGlyAspLysAlaTrpGlnTrpAsnAspAsnGlnMetTrpLeuPhe 625
 Db 2312 CCTAAATACAGCTCTGGAGATTAACCATATGAAATGGAAGCAATGAATGTACTCTGT 2371
 QY 625 ArgSerSerValAlaTrpAlaMetArgIleTrpPheLeuLysValLysAsnGlnMetIle 645
 Db 2372 CCGATATCTGTTGCTATATGCTATGAGGACGATCTTTTAAAGTAAAAATCAATGAT 2431
 QY 2432 TCTTTTGGGAGAGAGATGTGGAGTGCCTATTTGAACCAAGATCTCTTTAAATTT 2491
 QY 665 PheValIleThrAlaProLysAsnValSerAspIleIleProArgTrpGlnValGlnLysAla 685
 Db 2492 CTGTGCTACGCACTTAAATGCTGTATATCATCTCTAAGCACTGAAGTGAAGAAAGCC 2551
 QY 685 AlaLeuMetSerAspSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGln 705
 Db 2552 CATCAGAGATGTCCGGAGCGGATCATATGATGTTCCGCTGGAATGACAAAGCCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 Db 2612 GTTTCGGGGATACAGCCAAACTTGACCTCTTAACAGCCCTCT 2657
 RESULT 9
 US-10-123-904-71
 Sequence 71, Application US/10123904
 Publication No. US20030022328A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: DeForge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerlitsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ACIDS ENCODING THE SAME
 FILE REFERENCE: P3330R1054
 CURRENT APPLICATION NUMBER: US/10/123, 904

CURRENT FILING DATE: 2002-04-16
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 71
 LENGTH: 3732
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-123-904-71

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-123-904-71 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 |||
 Db 94 TCACCATGAGGAAAGCCCAAGACATTTTGGACAGTTTAAACCCAGAACCCAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThcGlnGluAsn 40
 |||
 Db 154 CTGTTCTATCAAAAGTTCACTGCTCTTGGAAATTATACACCATAATTACTGAAGAGAT 213
 QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGlnGlnSerThr 60
 |||
 Db 214 GTCCAAACATGAAATATCTGGGACAAATGTCCTGCTTTTAAAGAACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 |||
 Db 274 GTTGGCCAAATATCTCACTACAGAAATTCAGAAATCTCACAGTCAAGCTTCAGTGCAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 |||
 Db 334 GCTCTTCAACCAAAATGGCTTCAGTGCCTCTCAGAAAGACAGACAAAGCGTTGACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 |||
 Db 394 ATTTCTAAATACATGAGCACCACATCTACAGTACTGAGAAAGTTTGAACCCAGATATCA 453
 QY 121 GlnGluCysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 |||
 Db 454 CAAAGATGCTTATTTCTTGAACCAAGTTGAATGAATGAAGCAAAAGTTTGAACCTAC 513
 QY 141 AsnGluArgLeuTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPalaTyrPala 160
 |||
 Db 514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGGTGGGCAACAGCTGAGGCCA 573
 QY 161 LeuTyrGlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 |||
 Db 574 TTAATATGAAGATAGTGTCTTGAATAATGAGAGGCAAGGCAAAATCAATTAAGAGAC 633
 QY 181 TyrGlyAspTyrTyrArgGlyAspTyrGluValAsnGlyValAsnGlyTyrAspTyrSer 200
 |||
 Db 634 TATGGGATATATGAGAGAGAGACATAGAGTAATAGGAGTATGAGTATGAGTACAGGC 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluIleThrPheGlnGluIleLysProLeuTyrGlu 220
 |||
 Db 694 CGCGGCAAGTATGAGAGATGAGATACCTTGAAGATTAACCAATTAATATCA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 |||
 Db 754 CATCTTCATAGCTATAGTGAAGGCAAAAGTTGAATGATCCCTATCTATATCAGTCCA 813
 QY 241 IleGlyCysLeuProAlaHisLeuLeuLysMetTyrPalaArgPheTyrThrAsnLeu 260
 |||
 Db 814 ATTTGATGCTCCCTGCTCATTTGCTGTATAGTGGGTAGATTTTGAACAAATCTG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValIleAspAlaMetVal 280
 |||

Db 874 TACTCTTTGACAGTCCCTTTGGACAGAAACCAAAACATAGATCTTACTGATGCAATGCTG 933
 QY 281 AspGlnAlaTyrPalaGlnArgIlePheLysGlnIleGlyLysPhePheValSerVal 300
 |||
 Db 934 GACCAAGGCTTGGATGACAGAGAAATATCAAGAGGCGCAAGATCTTTGTATCTGTT 993
 QY 301 GlyLeuProAsnMetThrGlnGlyPheTyrPalaLysSerMetLeuThrAspProGlyLys 320
 |||
 Db 994 GGTCTTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTyrPalaLysGlyLysPheArgIle 340
 |||
 Db 1054 GTTCAGAAAGCAGTCCGATCCACAGCTTGGGAGCCCTGGGAGGCGACTTCAGATGC 1113
 QY 341 LeuMetCysThrLysValIleThrMetAspAspPheLeuThrAlaHisGlnMetGlyHis 360
 |||
 Db 1114 CTATGTGCACAAAGGAGGACATGACACTTCTGACAGCTCATCTAGTATGAGGAGAT 1173
 QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 |||
 Db 1174 ATCCAGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
 QY 381 GlyPheHisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 |||
 Db 1234 GGATTCATGAAAGCTGTTGGGAAATATGATGATGATGATGATGATGATGATGATGATGAT 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPro 420
 |||
 Db 1294 AAATCCATGCTCTCTGTCACCGATTTTCAAGAACATGAAACAAATTAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 |||
 Db 1354 CTGCTCAAAACAGCAGCAGATGTTGGAGCTCTGCACTTACTTACTTATGATGAGAG 1413
 QY 441 ThrArgThrMetValPheLysGlyGluIleProLysAspIleThrPheLysTyrPala 460
 |||
 Db 1414 TGGAGTGGAGGCTTTTAAAGGGGAAATCCCAAAACCAAGGAGTGAAGAAAGTGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValGlyProValProHisAspGluThrTyrCys 480
 |||
 Db 1474 GAGATGAAGGAGAGATAGTGGGCTGTGGAACTGTGCCCATGATGAAACTCTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
 |||
 Db 1534 GACCCGCACTCTGTTCCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleLysHisGlnGluPro 520
 |||
 Db 1594 ACCCTTACCAATTCAGATTTCAGAAAGCACTTGTGCACAGCTTAAACATGAAAGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 537
 |||
 Db 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGAGACAGAAACTGTT -GTAAAGAAAT 1712
 QY 537 ----- 537
 Db 1713 ACCTCAAAATGTTGAACCTCTCTAGTATTCAGTATTCATTCATTCAGTCCAGCTAGGTTTG 1772
 QY 537 ----- 537
 Db 1773 TATTTGATTTCTTTGTTCTTAAAAAGAAATTTTATGCGCTCAAAATGCTCTCATTTACAA 1832
 QY 537 ----- 537
 Db 1833 ACCAAACATTTAATTTGTGTGAGACAGAAAGCACTAGACATACAAACATTTGGGTGGCCA 1892
 QY 537 ----- 537
 Db 1893 CCTCTTTTCTCCATCATTAACAGAGCCCTCTCTCTGTATTTGAAGAAAGAGCG 1952
 QY 537 ----- 537
 Db 1953 GTTTAGGCTGGAATATATCTGTTAATATGCAATCTTTCTTATCTGCGCAAGCAAAATTT 2012

281 ASPGlnAlaTrpAspAlaGlnArgIlePheIysGlnAlaGlnIysPhePheValSerVal 300
 934 GACCAGGCGTGGGATGCCACAGAGATATTCAGAGAGGCGAGAGATTCTTGTATCTGTT 993
 301 GlyLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 320
 994 GGTCTCCATATATGATCAACAGAGATTCGGGAAAAATTCATCCATACGAGACCCAGGAAT 1053
 321 ValGlnIysAlaValAlaCysHisProThrAlaTrpAsnLeuGlyIysGlyAspPheArgIle 340
 1054 GTTCAGAAAGAGCTGGCATCCACACCTGGGAGACCTGGGAGAGGCGGACTTCAGAGATC 1113
 341 LeuMetCysThrIysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyHis 360
 1114 CTATGTCACAAAGGTGACATGACATGACATCTTCTGACAGCTCATGAGATGGGCGAT 1173
 361 IleGlnIysAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGln 380
 1174 ATCCAGTATGATATGGCATATGTCGACCAACCTTTCTGTAAGAAATGGAGCTAATGAA 1233
 381 GlyPheHisGlnAlaValAlaGlyGlnIleMetSerLeuSerAlaAlaThrProIysHisLeu 400
 1234 GGATTCATGAGAGCTGGTGGGAAATCATGTCACCTTCTGCGAGCCACCACTAAGCATTTA 1293
 401 LysSerIleGlyLeuLeuSerProAspPheGlnGlnAsnGlnIleThrGlnIleAsnPhe 420
 1294 AAATCCATGCTCTCTCTGTCACCGATTTTCAGAAACATGAAACAGAAATTAACCTTC 1353
 421 LeuLeuIysGlnAlaLeuThrIleValAlaGlyThrLeuProPheThrIleMetLeuGlnIys 440
 1354 CTGCTCAACACAGACACACAGATGTTGGGACCTGCGCATTTCTTCAATGTTAGAGAAC 1413
 441 TrpArgTrpMetValPheIysGlnIleProIysAsnGlnIleMetIysIleTrpTrp 460
 1414 TGGAGGGGAGTGGCTTTAAAGGGGAAATCCCAAAAGACAGGAGGAAAGAGTGGTGG 1473
 461 GluMetLysArgGlnIleValAlaGlyValAlaGlnProValProHisAspGlnIleThrCys 480
 1474 GAGATGAGAGGAGAGATAGTGGGGTGGTGGAACTTGGCCCATGATGAAACATCTGCT 1533
 481 AspProLaserLeuPheHisValSerAsnAspIleSerPheIleArgIleTrpTrpArg 500
 1534 GACCCCGCATCTGTTCCATGTTTCATGATGATGATGATGATGATGATGATGATGATGAT 1593
 501 ThrLeuTrpGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaIysHisGlnIlePro 520
 1594 ACCCTTACCAATTCAGTTTCAAGAGACCTTGTCAACGACCTAAGCTGAAGGCCCT 1653
 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyGlnIysLeuPhe 537
 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTT-GTAAAGAAAT 1712
 537 537
 1713 ACCTCAAAATGTTGAACCTCTCCTAGTATTCATGATTCATGCTAGGTTTG 1772
 537 537
 1773 TATTGATTTCTTTGTTCTAAAGAAATTTTATGCGCTCAAAATGCTCATTTTCAAA 1832
 537 537
 1833 ACCAAACATTTAATTTGTGTCAGACAGAGACCTACACATACAAACATTTGGGTGGCCA 1892
 537 537
 1893 CCTCTTTTCTCCTATCATACACAGCCCTCTCTCTGTAATTTGAAGAAAGAGCG 1952
 537 537
 1953 GTTTAGGTTGGAATATATCTGTTAATATGCAATCTTTTCTATCTGCGAGAAAGAAATTT 2012
 537 537

2013 AGCCAGCTCAAGAGAGAAACCATAGATCATAGATGTAATATATGATCTGAGAAC 2072
 538 -----AsnMetLeuArgLeuGlyIysSer 545
 2073 CTTCAAAAGGCCCTGAACCCCTTTTGTGTAGCAAAATATGCTGAGCGCTGGAATATCA 2132
 546 GluProTrpThrLeuAlaLeuGlnAsnValAlaGlyValIysAsnMetSerValArgPro 565
 2133 GAACCGTGGACCTTACATGATGGAATATGTTGTAGAGCC-AGGAACTGAAATGTAGGCCA 2191
 566 LeuLeuAsnTrpPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 2192 CTGCTCAACATCACTTGTAGCCCTTATTTACCTGGCTGAAGACCAAGCAAGAAATCTTTT 2251
 586 ValGlyTrpSerThrAspTrpSerProIysAlaAsp-GlnSerIleLysValArgIleSe 605
 2252 GTGGGATGGAGTACGAGCTGAGGTCCATATGACAGACCAACCAACATCAAGATGATAG 2311
 605 IleLysSerAlaLeuGlyAspLysAlaTrpGlnTrpAsnAspAsnGlnMetTrpLeuPhe 625
 2312 CCTAAATACAGCTCTTGAGATTAACATATGATGAAACGCAATGAAATGTACCTGT 2371
 625 eArgSerSerValAlaTrpAlaMetArgIleTrpPheLeuIysValLysAsnGlnMetIle 645
 2372 CCGATCATCTGTTGATATGATGATGAGCAGTACTTTTAAAGTAAATCAAGATGAT 2431
 645 eLeuPheGlnGlnGlnAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 2432 TCTTTTGGGAGAGAGATGCTGAGGTGCTAATTTGAAACCAAGAAATCTCTTTAAATTT 2491
 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGlnValAlaGlnIysAl 685
 2492 CTTTGTCACTGCACCTAAATATGTTGTGATATCATTTCTTGAAGCTGAAGTTGAAAGGC 2551
 685 aLeuArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAsnSerLeuGln 705
 2552 CATCAGAGATGCTCCGAGCGGATCAATGATGATGATGATGATGATGATGATGATGATGAT 2611
 705 uPheLeuGlnIleGlnProThrLeuGlnIleProProAsnGlnProPro 720
 2612 GTTTCGGGAGATACAGCAACACTTGACCTCTCAACAGGCCCT 2657

RESULT 11
 US-10-175-746-71
 ; Sequence 71, Application US/10175746
 ; Publication No. US20030027270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C353
 ; CURRENT APPLICATION NUMBER: US/10/175,746
 ; PRIORITY FILING DATE: 2002-06-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71

```

; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-746-71

Alignment Scores:
Pred. No.: 0          Length: 3732
Score: 3720.00        Matches: 717
Percent Similarity: 83.88%      Conservative: 1
Best Local Similarity: 83.76%    Mismatches: 2
Query Match: 96.15%             Indels: 138
DB:                               Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-175-746-71 (1-3732)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db 94 TCACACATGAGGAGCAAGCCAAAGACATTTTGGACAACTTTAACACAGAACCCGAGAGC 153
QY 21 LeuPheTyGlnSerSerLeuAlaSerTPAsnTYrAsnThrAsnIleThrGluGluAsn 40
Db 154 CTGTTCTATCAAAAGTCACTCTCTCTGGAATTTAACCCAAATATTTACTAGAGAAAT 213
QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db 214 GTCCAAAACATGATATATGCTGGGACAAATGCTGCTTTTAAAGAGACAGTCCACA 273
QY 61 LeuAlaGlnMetTyRProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db 274 CTGGCCAAATGATACCTACACAAAGAAATTCAGAAATCTCACAGTCAAGCTCAGCTG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
Db 334 GCTCTTCACCAAAATGGGCTTCAGGCTCTCAGAGAACAGAACGAGGTTGAACACA 393
QY 101 IleLeuAsnThrMetSerThrIleTyRSerThrGlyLysValCysAsnProAspAsnPro 120
Db 394 ATCTCAATATACATAGACACACATCTACAGTCTGAGAAAGTTTGTAAACCACTAATCCA 453
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyR 140
Db 454 CAAGATGCTTATCTTACCTGACACAGGTTGAATGAATATGCAAAACAGTTTACACTAC 513
QY 141 AsnGluArgLeuThrAlaTrpGluSerTPArgSerGluValGlyLysGlnLeuAspPro 160
Db 514 AAGGAGAGGCTCGGGCTGGGAAAGCTGAGATCTGGGCAAGCAGCTGAGGCCA 573
QY 161 LeuTyRGlulGluTyRValValLeuLysAsnGluMetAlaArgAlaAsnHisTyRGlulAsp 180
Db 574 TTATATGAGAGATATGCTGCTTGAATAATGAGATGGCAAGACAAATCATTTATGAGAC 633
QY 181 TyrGlyAspTyRTPArgGlyAspTyRGlulValAsnGluValAspGlyTyRAspTyRSer 200
Db 634 TATGGGGAATATTTGAGAGAGAGACTATGAATAAATGGGGTATGATGCTATGCTACAG 693
QY 201 ArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyRGlul 220
Db 694 CGGGCCAGTGGTAAAGATGTGGAACATACCTTTGAAGAGATTTAAACCATTTATGTGAA 753
QY 221 HisLeuHisAlaTyRValArgAlaLysLeuMetAsnAlaTyRProSerTyRHisLeuPro 240
Db 754 CATCTTCATGCCCTATGTGAGGAGCAAGTTGATGAATGCCCTATCTCTATATCATCTCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyRValArgPheTyRHisLeu 260
Db 814 ATTGAGATGCTCCCTGCTCATTTGCTGGATATGTGGGGTATGATTTTGGACAAATCTG 873
QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
Db 874 TACTCTTTGACATCTCCCTTTGGACAGAAACCAACATAGATGTCTGTAAGCAATGGTG 933
QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300

```

```

Db 934 GACCAGGCTCGGATGACACAGAAATATTCAGGAGGCCGAGAAAGTCTTTGTATCTGTT 993
QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
Db 994 GGTCTTCCTATATGACTCAAGAGATTCGGGAAATTCATGATCAAGGACCCGAGAAAT 1053
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
Db 1054 GTTCAGAAAGAGCTCGCATCCACACACCTTGGGACCTGGGGAGGCGACCTTCAGATC 1113
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisLysGluMetGlyHis 360
Db 1114 CTATATGTCACAAAGTGCACAAATGAGACACTCTCTGACAGCTCATCATGATGGGCAT 1173
QY 361 IleGlnTyRAspMetAlaTyRAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1174 ATCCAGATGTATGATGTCATATGCTGACAAACCTTTCTGTAAAGAAATGAGCAATAGAA 1233
QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1234 GGATTCATGAGAGCTGTGGGAAATCATGTCTCTCTGACGCCACACCTTAACCAATTA 1293
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
Db 1294 AATTCATTTGGCTCTCTCTGACCCGATTTTCAAGACATGAAACAGAAATTAACCTTC 1353
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyLysLeuProPheThrTyRMetLeuGluLys 440
Db 1354 CTGCTCAAAACAGCACTCAGCATTTGTTGGACTGTCGCAATTTACTTACATGTAGAGAA 1413
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysTyRTrp 460
Db 1414 TGGAGTGGAATGCTTTAAAGGGAAATTCACAAAGCAAGTGGAATGAATAATGGTGCG 1473
QY 461 GluMetLysArgGluIleValGlyValGlyAlaGluProValProHisAspGluThrTyR 480
Db 1474 GAGATGAGCGAGAGATATGTTGGGTGGTGAACCTGCGCCCATGATGAACATATCTGT 1533
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyRSerPheIleArgTyRThrArg 500
Db 1534 GACCCCGCATCTCTGTCATGTTCTGATGATTAACATCATTCATTCATATACACAAG 1593
QY 501 ThrLeuTyRGlulPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGluPro 520
Db 1594 ACCCTTTACCAATTCAGATTCAGAAACACCTTGTCAAGACGCTAAACATGAAGGCCCT 1653
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
Db 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAGCTGACAGAAACGTGT-GTAAAGAAAT 1712
QY 537 ----- 537
Db 1713 ACCCTAAATGTGAACCTCTCTAGTATTCAGTATTAATTCATGCTAGGTTTG 1772
QY 537 ----- 537
Db 1773 TATTTGATTTCTTTGTCTTAAAAAGAAATTTATGGCTCAAAATGTCTCATTTACAA 1832
QY 537 ----- 537
Db 1833 ACCCAAACTTTAATTTGTGGTCAGACAGAAACCTGACATACCAACAAATTTGGGTGGCA 1892
QY 537 ----- 537
Db 1893 CCTCTTTCTCCCTATCATATACAGCCCTCTCTCTGGTAATTTGAGAAAGAGAGCG 1952
QY 537 ----- 537
Db 1953 GTTTAGGGTGAATATATCTGTTAATATGATCTTTCTTATCTCCAGAAACAAATTT 2012
QY 537 ----- 537
Db 2013 ACCCAAGTCAAGAGAAAGAAACCATAGATCATAGATGTAAATATATGTACATCTGGAAC 2072

```


QY 538 -----AsnMetLeuArgLeuGlyLysSer 545
 DB 2073 CCTCAAAAGCCCTGACCCCTTTTGTGTACCAATGCTGAGGCTTGGAATATCA 2132
 QY 546 GIUPROTPTPTHTLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565
 DB 2133 GAAACCTGGACCTTGACATGGAATATGTTGAGAGC-AAGAACATGAATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTyrPheLeuLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCAACTACTTGTAGCCCTTATTTACCTGCGTGAAGAACCAACAAGATTTCTTT 2251
 QY 586 ValGlyTyrPserThrAspTyrPserProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 DB 2252 GTGGATGAGATGACCTGAGCTGAGTCCATATGAGACCCAAAGCAATCAAGAGCATAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAspAsnGlnMetTyrLeuPhe 625
 DB 2312 CCTAAATACAGCTCTTGAGATGAAGCATATGATGAACGCAATGAATGTAAGTACTGTT 2371
 QY 635 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
 DB 2372 CCGATCATCTGTTCATATGCTATGAGCAGTACTTTTAAAGTAAATAATGAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGAGGAGATGTCGAGTGGCTAATTTGAACCAAGAAATCTCTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlyLysAl 685
 DB 2492 CTTTCTACATGACTTAAATATGTCTGATATCTTTCTTGAACTGAAAGTTGAAGAGC 2551
 QY 685 allLeuMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 DB 2552 CATGAGATGTCCCGAGACCCATCAATGATGATGCTTCGCTGATGCAACAGCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGAGATACAGCACACTTGACCTCTCAACAGCCCTT 2657
 RESULT 12
 US-10-176-918-71
 ; Sequence 71, Application US/10176918
 ; Publication No. US20030027275A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3330R1C382
 CURRENT APPLICATION NUMBER: US/10/176, 918
 PRIORITY FILING DATE: 2002-06-20
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 71
 LENGTH: 3732
 TYPE: DNA

; ORGANISM: Homo Sapien
 US-10-176-918-71
 Alignment Scores:
 Pred. No.: 0
 Score: 3720.00
 Percent Similarity: 83.88%
 Best Local Similarity: 83.76%
 Query Match: 96.15%
 DB: 9
 Indels: 138
 Gaps: 1
 US-09-978-385-2_COPY_19_738 (1-720) x US-10-176-918-71 (1-3732)
 QY 1 SerThrIleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
 DB 94 TCCACCATTTGAGGACAGGCGCAAGACATTTTGGACAGATTACCAACGAGCCGAGAC 153
 QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnTyrAsnIleThrGluGluAsn 40
 DB 154 CTGTTCTATCAAGTTCACTTCTCTGGAATTAACCAATATTAAGTACTGAAAGAT 213
 QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrPserAlaPheLeuLysGluGlnSerThr 60
 DB 214 GTCCAAATCATGATTAATGCTGGGACAAATGATGCTCTTTTAAAGAACAGTCCACA 273
 QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
 DB 274 CTGGCCAAATGATACACTACAGAAATTCAGAAATCTCAAGTCAAGCTTCAGCTGAG 333
 QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
 DB 334 GCTCTTACGAAATGGGTCTTCAGTGCCTCAGAAAGACAGCAAGGTTGAACACA 393
 QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
 DB 394 ATTCTAATCATGATGAGCACCACATCTACAGTCTGGAAGTTGTGAACCAATATCA 453
 QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
 DB 454 CAAGAAATGCTTATTAATTGAAACAGGTTGAATGAATGAATGGAACAAAGTTAGACTAC 513
 QY 141 AsnGluArgLeuThrPheLysTyrPheGlnSerTyrPheArgLysGluValGlyLeuArgPro 160
 DB 514 AATGAGAGGCTCTGGGCTGGGAAAGCTGAGATCTGAGAGTGGCAAGCAGCTGAGCCA 573
 QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
 DB 574 TTATATGAGAGATATGCTGCTGGAATAATGAGATGCAAGCAATATTAATGAGAC 633
 QY 181 TyrGlyAspTyrTyrPheArgLysPheTyrGluValAsnGlyValAspGlyTyrAspTyrSer 200
 DB 634 TATGGGAGATTAATGAGAGAGACTATGAATGAATGAGGATGATGATGATGATGATGATG 693
 QY 201 ArgGlyGlnLeuIleGluAspValGluHisTyrPheGluGluIleLysProLeuTyrGlu 220
 DB 694 CGCGCCAGTGTATTAAGATGAGAACATACCTTTGAAGATTAACCATTAATATGAA 753
 QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
 DB 754 CATCTCATGCGATATGAGGCAAGTGAATGCTTATCCATTCATATGATGATGATGATGATG 813
 QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrPheTyrPheThrAsnLeu 260
 DB 814 ATTGATGCTCTGCTGCTATTTGCTGATGATGATGATGATGATGATGATGATGATGATG 873
 QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
 DB 874 TACTCTTACAGTTCTCTTGGACGAAACCAATCAATGATGATGATGATGATGATGATGATG 933
 QY 281 AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
 DB 934 GACACAGGCTGGGATGACAGAGAAATATCAAGAGAGGCGGAGAAATCTTTGATATCTGT 993

QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn 320
 DB 994 GGGCTTCCTCAATATGACTCAAGATTCGTGGAAATTCGTAACGAGACCCAGGAAT 1053
 QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
 DB 1054 GTTCAGAAAGACATCTGCGCATCCACAGCTTGGAGACGTGGAGAGGCGACTTCAGATC 1113
 QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisHisGlnMetGlyHis 360
 DB 1114 CTATGCGCAAGGAGCAATGAGACAGCTTCGACAGCTCATGATGAGTGGGCAAT 1173
 QY 361 IleGlnTrpAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATGAGCAATGCTGCACAACTTTCTGCTAAGAAATGAGCTAATGAA 1233
 QY 381 GlyPheHisGlnAlaValAlaGlyLuiLemSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GGAATTCATGAACTTTGGGGAATCATGTCCTTTCGACGCCACACTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1294 AAATCCATGTCCTCTGTCACCCGATTTTCAAGAGACATGAAACAGAAATAACTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGlyLys 440
 DB 1354 CTGCTCAACCAACACCTCAGATTTGGGACCTCCATTTACTTACATGTTAGGAAG 1413
 QY 441 TrpArgTrpMetValPheLysGlyLuiLemProLysAspGlnTrpMetLysLysTrp 460
 DB 1414 TGGAGGTGATGCTCTTAAAGGGGAATTTCCAAAGACACGATGATGAAAAAGTGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValAlaGluProValProHisAspGluThrIleCys 480
 DB 1474 GAAATGAAGGAGAGATAGTTGGGGGTGGGAACCTGCCCCATGATGAACATPACTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheLeuArgIleGlyThrArg 500
 DB 1534 GACCCCGCATCTGTTCCATGTTCTGATTTACTCATTTCTGAAATTAACAAAG 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyPro 520
 DB 1594 ACCCTTACCATTTCCAGTTTCAAGAAAGCATTTGTCAGACAGCTAAACATGAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAAAGTGCAGAAAGTGT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCTCAAAATGTTGAACCTCTCCTAGATTCAAGTATTCTCATTTCCAGCTAGGTTTG 1772
 QY 537 ----- 537
 DB 1773 TATTGATTTCTTGTCTAATAAAGAAATTTTATGGCTCAAAATGTCCTCATTTACA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTAAATTTGTGTCAGACAGAACCTAGACCATCAACAATTTGGTGGGCCA 1892
 QY 537 ----- 537
 DB 1893 CCTCTTTTCCCTTATCATTACTACAGCCCTCTCTCTGTAATTGGAAGAAAGAGCG 1952
 QY 537 ----- 537
 DB 1953 GTTAGGGTGAATATATCTGTTAATATGATCTTTCTTATCTGCCAGAAACAATTT 2012
 QY 537 ----- 537
 DB 2013 AGCCAAGTCAAAAGAGAAACCATAGATCATAGATGTAATATATGTACATCTGGAACC 2072
 QY 538 ----- 545
 AsnMetLeuArgLeuGlyLysSer

DB 2073 CCTCAAAAGCCCTGAAACCCCTTTTGTGTAGCAATATGAGGCTTGAATAATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAspValArgPro 565
 DB 2133 GAACCTGACCTTGCATTTGGAATATTTGTAGAGC- AAGACATGAATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCACTACTTTGACCCCTTATTTACCTGGCTGGAAGACCAACAACAATTTCTTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 DB 2252 GTGGGATGAGTACCCGACTGTGAGTCATATGCAGACCCCAAGACATCAAAAGTGAAGTAA 2311
 QY 605 rLeuLysSerAlaLeuGluAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPhe 625
 DB 2312 CCTAAATACAGCTCTTGGAGATTAAGCATATGATGGAACACATGAATATGACCTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetLil 645
 DB 2372 CCGATCATCTGTTGCATATGCTATGAGCAGACTTTTAAAGTAAATTCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGGAGGAGATGTGCGAGTGGCTAATTGAAACCAAGAAATCTCTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleLeuProArgThrGluValGlyLysAl 685
 DB 2492 CTTTCTACCTGCACCTTAAATATGTTCTGATATCTTCTTGAACCTGAAGTTGAAAGGC 2551
 QY 685 alleaRgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 DB 2552 CATCAGATGTCGCCGAGCCGATCATATGATGCTTCCGCTGATGACACACACCTTGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGGATACAGCCACACTTGACCTCAACAGGCCCT 2657
 DB
 RESULT 13
 US-10-176-921-71
 ; Sequence 71, Application US/10176921
 ; Publication NO. US20030027276A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroft, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C288
 ; CURRENT APPLICATION NUMBER: US/10/176,921
 ; PRIORITY FILING DATE: 2002-06-20
 ; PRIORITY FILING DATE: 2002-06-20
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 71
 ; LENGTH: 3732
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-176-921-71

Qy		546	GILPrrPhrThrlleuAlaleuGluAsnValIglValalysAsnMetAsnValArPro	565
Dd		2133	GAACCTGGACCCTAGCATTTGGAAAAATCTTGTAGAGAC - AAGAACAATGAATGTAAAGGCCA	2191
Qy		566	LeuluAsnTyrrPhegiuProleuPheThrTripleuLysaspGlnasnlYasnSerPhe	585
Dd		2192	CTGCCTCAACACTTTTAGCCCTTATTATTAAGTGCGGTGAAGACCGAACCAAGAAATTCCTTT	2251
Qy		586	VaIGlYrrSerTrhrAsPTPrSerProTYrAlaasp-GInserTlleYsnValArgIese	605
Dd		2252	GTGGAGATGGATGACCGACTGGATGCTCAATATGCAAGACCCAAGCATCAAGTGAAGATPAG	2311
Qy		605	rLeuLysSerAlaleuGlyAspLysAlaTyrgluTrpAsnAspAsngluMetYrleuPh	625
Dd		2312	CCATAAATVACGCTCTTGAGATTAAGCATATGATGATGACAGACATGAAATGTACCTGTT	2371
Qy		625	eArGSeSseValAlaTyrrAlaMetArgGlnTYrPheleuLysValLysAsngluMetII	645
Dd		2372	CCGATVCTCCTGGTCATATGCTATGAGGACGACTCTTTTAAAATPAAAAAATVACAGATAT	2431
Qy		645	eLeuPhgIGluGluAspValArgValAlaAsnLeuLysProArgIlesePheAsnph	665
Dd		2432	TCTTTTTGGGAGAGAGATGTGCAGATGGCTAAATTTGAACCAAGAAATCTCTTAATTT	2491
Qy		665	ePheValThrlaProLysAsnValSerAspIleIeProArgThrGIuValGluLysAI	685
Dd		2492	CTTGTGTACGTGCACCTTAATAATGTCTGATATATTCCTGAACCTGAAGTTGAAAAAGCC	2551
Qy		685	aLeArMeSerArgSerArgIleAsnAspAlaPheArgleuAsnAspAsnSerleuGI	705
Dd		2552	CATAGAGATGCCCCGGAGCCGTATCAATGATGCTTTCGTGTGATGATGCAACAGCTAGA	2611
Qy		705	uPheLeuGIylleGlnProThrlleuGlyProProAsngluInProPro	720
Dd		2612	GTTCCTGGGGATACAGCCAACACTTGGACCTCCTAACCAGACGCCCT	2657
RESULT 14				
US-10-137-865-71				
; Sequence 71, Application US/10137865				
; Publication NO. US20030032155A1				
; GENERAL INFORMATION:				
APPLICANT: Baker, Kevin P.				
APPLICANT: Beresini, Maureen				
APPLICANT: Deforge, Laura				
APPLICANT: Desnoyers, Luc				
APPLICANT: Filvaroff, Ellen				
APPLICANT: Gao, Wei-Qiang				
APPLICANT: Gerritsen, Mary E.				
APPLICANT: Goddard, Audrey				
APPLICANT: Godowski, Paul J.				
APPLICANT: Gurney, Austin L.				
APPLICANT: Sherwood, Steven				
APPLICANT: Smith, Victoria				
APPLICANT: Stewart, Timothy A.				
APPLICANT: Tumas, Daniel				
APPLICANT: Watanabe, Colin K				
APPLICANT: Wood, William				
APPLICANT: Zhang, Zemin				
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC				
TITLE OF INVENTION: ACIDS ENCODING THE SAME				
FILE REFERENCE: P3330R1C154				
CURRENT APPLICATION NUMBER: US/10/137,865				
CURRENT FILING DATE: 2002-05-03				
Prior Application removed - See Palm or File Wrapper				
NUMBER OF SEQ ID NOS: 550				
SEQ ID NO 71				
LENGTH: 3732				
TYPE: DNA				
ORGANISM: Homo Sapien				
US-10-137-865-71				

Pred. No.:	0	3720.00	Length:	3732
Percent Similarity:	83.88%		Matches:	717
Best Local Similarity:	83.76%		Conservative:	1
Query Match:	96.15%		Mismatches:	2
DB:	9		Indels:	138
			Gaps:	1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-137-865-71 (1-3732)

QY	1	SerThrIleGIuGIuGlnAlaIalysThrPheLeuAspLysPheAsnHisGIuAlaGIuAsp	20
DB	94	TCCACCAATGAGGAACAGGCCCAACACATTTTGGACAACTTATACCAACGAAAGCCAAAGAC	153
QY	21	LeuPheTyrGlnSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGluAsn	40
DB	154	CTGTCTATCAAGATTCCTTGCTTCTTGGAAATTATACCAACCATATTACTGAAGCAAT	213
QY	41	ValGlnAsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGlnGlnSerThr	60
DB	214	GTCACAAACATGAATATATGCTGGGACAAATGCTCCCTTTTAAAGAAACAGTCCACA	273
QY	61	LeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln	80
DB	274	CTTGCCCAAAATGATTCACACTCAAGAAATTCAGATCTCACAGCAAGCTTCAGCTGCAG	333
QY	81	AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysTyrGluAsnThr	100
DB	334	GCTCTTCAGCAAAATGGGCTTCAGCTCTCCACAAACAAAGACAAACGGTTGAACACA	393
QY	101	IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro	120
DB	394	ATTCTAAATACAAATGAGACCATCTACAGTACTGGAAGAGTTGTAAACCATTAATCCA	453
QY	121	GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140
DB	454	CAGAGATCTTATTAATCTTGAACACAGAGTTGTATAAATATGGCAAAACAGTTTAGACTAC	513
QY	141	AsnGluTyrGlnLeuThrPalaTyrGlnSerTyrPargSerGluValGlyLysGlnLeuAspPro	160
DB	514	AATGACAGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTGGGCAAGCAGCTGAGGCCA	573
QY	161	LeuTyrGlnGluTyrValValIleuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180
DB	574	TTATATGAAGAGTATGCTGCTTCAAAAATGATGCGCAGACCAAAATCATTTATGAGAGC	633
QY	181	TyrGlyAspTyrTyrPargGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer	200
DB	634	TATGGGGATTATTTGGAGAGAGACTATGAATGAATGGGGTAGATGGCTATGACTACAGC	693
QY	201	ArgGlyGlnLeuIleGlyAspValGluIsthrPheGlnGluIleLysProLeuTyrGlu	220
DB	694	CGCGGCCAGTATGATGAAGATGTGGAACATACCTTTGAAGAGATTAAACCATTTATAGAA	753
QY	221	HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro	240
DB	754	CATCTTCATGCCTATGATGAGGCAAAAGTTGATGATGCGTATCCCTTCATATACGTCCA	813
QY	241	IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTyrGlyAspPheTyrPheAsnLeu	260
DB	814	ATTGGATGGCTCCCGCTCATTTGCTTGATATATGTGGGTATGATTTGGACAAATCTG	873
QY	261	TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal	280
DB	874	TACTCTTGACAGTCCCTTGTTGAGACAGAAACCAACATAGATGTACGATGCAATGCTG	933
QY	281	AspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGlyLysPhePheValSerVal	300
DB	934	GACACAGGCTCGGAGATGACAGAAATATATTCAGAGGCGCAGAAAGTTCTTGTATCGTT	993
QY	301	GlyLeuProAsnMetThrGlnGlyPheTyrPgiLysAsnSerMetLeuThrAspProGlyAsn	320
DB	994	GGCTCTCTAAATATGACTCAGAGATTTCTGGGAAATTCATGCTTACGAGACCCAGGAAT	1053

```

QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTle 340
Db 1054 GTTCAGAAAGACAGTCTGCCATCCACAGCTTGGGAGGAGGCGACCTTTCAGAGC 1113
QY 341 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 360
Db 1114 CTATGTGACAAAGGTGACAAATGGACAGCTTCGACAGCCTCATCATGATGGGCGCT 1173
QY 361 IleGlnTrpAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
Db 1174 ATCCAGTATGATATGGCATATGCTGCACAACTTTCCTCTAAGAAATGAGACCTATGAA 1233
QY 381 GlyPheHisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
Db 1234 GGATTCAGAAAGCTGTGGGAAATCATGCTACTTTCGACGACCACTTAAGCATTTA 1293
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
Db 1294 AATATCCATGCTCTTCGTCACCCGATTTTTCAGAAAGCAATGAAACAAATAAATTC 1353
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
Db 1354 CTGCTCAAGCAAGCACTCAGATTTGGGAGCTGCGCATTTACTTACATGTTAGAGAG 1413
QY 441 TrpArgTrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrp 460
Db 1414 TGGAGTGGATGCTTTAAAGGGAAATTCACAAAGCCAGTGAATAAAGTGGGG 1473
QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
Db 1474 GAGTGAAGACGAGATAGTGGGGTGGGAGCTGCCCATGATGAAACATACGT 1533
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db 1534 GACCCCGATCTCTTTCATGTTTCTGATATGATTCATTCATTCGATTAACAGAG 1593
QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluPro 520
Db 1594 ACCCTTTCACCAATTCAGATTCAAGAGACACTTTGTCAAGAGCTAAACATGAGAGCCCT 1653
QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPhe----- 537
Db 1654 CTGCACAAATGTGACATCTCAAACTCAAGAGCTGACAGAAACGTGT- GTAAGAAAT 1712
QY 537 ----- 537
Db 1713 ACCTCAAAATGTGACCTCTCTAGATTCTAGATTACTTCATTCATTCAGCTAGGTTTG 1772
QY 537 ----- 537
Db 1773 TATTTGATTTCTTGTCTAAAGAAATTTATGGCTCAAAATGCTCATTTACAA 1832
QY 537 ----- 537
Db 1833 ACCAACAATTAATTTGTGTGACAGACAGAACTAGACATACAAACATTTGGGTGGCCA 1892
QY 537 ----- 537
Db 1893 CCTCTTTTCCCATATACATACAGACCTCTCTCTCTGTAATTGGAAAGAAAGCG 1952
QY 537 ----- 537
Db 1953 GTTTAGGTGGAATATATCTGTAATATGATCTTTTCTTATCTGCAGAAACAATTT 2012
QY 537 ----- 537
Db 2013 AGCCAACTCAAGAGAAAGAAACATAGATGTAATATATATGATCATCTGGAACC 2072
QY 538 ----- 538
Db 2073 CCTCAAAAGGCGCTGAGACCCCTTTTGTGTAGCAATATGCTGAGGCTTGAAATATCA 2132
QY 546 GluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565

```

```

Db 2133 GAACCTGAGACCTAGCATTTGAAATATGTTAGAGC -AAGAACATGATGTAAGCCA 2191
QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysSerPhe 585
Db 2192 CTGCTCAACTATCTTTGAGCCCTTATTTACCTGGCTGAAAGACAGAAAGAAATTCCTTT 2251
QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
Db 2252 GTGGAGTGGATACCGATGAGCTGATGCAATGCAAGCCAAAGCATCAAGTGAAGTAAAG 2311
QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
Db 2312 CCTAAATCAGCTCTTGGAGATTAAGCATGATGATGAGGAGTACTTTTAAAGTAAATCAGATCAT 2371
QY 625 eArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
Db 2372 CCGATCATCTGTTGCAATATGCTATGAGGAGTACTTTTAAAGTAAATCAGATCAT 2431
QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
Db 2432 TCTTTTGGGAGAGAGATGTGCGAGTGTATTTGAACCAAGATCTCTTTAATTT 2491
QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
Db 2492 CTTTGTACACGACCTTAAATATGCTGATATCATTCATTCAGAACTGAAGTTGAAAGGC 2551
QY 685 aileArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspSerLeuG 705
Db 2552 CATCAGATGTCCCGGAGCCGCTTCAATGATGCTTCCGCTGAAGACAGAGCTTGA 2611
QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
Db 2612 GTTCTGGGAGATACAGCAACACTTGAGCTCTTAACAGCCCTT 2657

RESULT 15
US-10-140-474-71
: Sequence 71, Application US/10140474
: Publication No. US20030032156A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3330R1C162
: CURRENT APPLICATION NUMBER: US/10/140,474
: CURRENT FILING DATE: 2002-05-06
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 71
: LENGTH: 3732
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-140-474-71

Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 3720.00 Matches: 717

```


QY 341 LeuMetCysThrIysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHis 360
 DB 1114 CTTATGTGACAAAGGAGACATGACACTCTCTGACAGCTCTCACTGAGATGGGACAT 1173
 QY 361 IleGlnIleAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATATGGCATATGCTGCACAAACCTTTCTGCTAAGAAATGAGAGTAAAG 1233
 QY 381 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProIleHisLeu 400
 DB 1234 GGATTCATGAGAGCTGTGGGAAATCATGCTCTCTGACAGCCACCACTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAsnGluThrGluIleAsnPhe 420
 DB 1294 AAATCCATGTGCTTCTGTGTCACCGATTTTCAGAAACATGAAACAGAAATTAACCTTC 1353
 QY 421 LeuLeuIleGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLys 440
 DB 1354 CTGCTCAACAGACAGACATGCTGGGACTCTGCCATTTACTTACATGTTAGAGAG 1413
 QY 441 TPATGTPMetValPheLysGlyGluIleProLysAspGlnIleProMetLysLysTyrTrp 460
 DB 1414 TGGAGGTGAGGTCTTTAAAGGGGAAATTCACCAAGACAGTGGATGAAAGAGTGGTG 1473
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHisAspGluThrTyrCys 480
 DB 1474 GAGATGAGAGCGAGATAGTGGGGTGGGACCTGGCCCGCATGAGAAACATCTGT 1533
 QY 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1534 GACCCCGCATCTGTTCCATGTTCTGATGATATACATCATTCATATTAACAAG 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuGluGlnAlaIleLysIleGlyPro 520
 DB 1594 ACCCTTACCAATTCAGTTCAGTTCAGAGACCTTGTCTGACAGCTTAAACATGAGGCCCT 1653
 QY 521 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTCTACAGAACTGGACAGAACTGTT- GTAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACCTCAAAATGTGACACTCTCTAGTATTCATGATTCATTCATGCTAGGTTTG 1772
 QY 537 ----- 537
 DB 1773 TATTTGATTTCTTTGTTCTAAAGAAATTTATGAGCTCAAAATGCTCATTTACAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTTAATTTGTGTGACAGACAGAACTTACACCATCAACATTTGGGTGGCCA 1892
 QY 537 ----- 537
 DB 1893 CCTCTTTCTCCATATCACTACAGCCCTCTCTCTGTAATTTGGAAGAAAGAGCG 1952
 QY 537 ----- 537
 DB 1953 GTTTAGGCTGAATATATCTGTATATGATCATCTCTTTCTTATCTGCCAGAACCAATTT 2012
 QY 537 ----- 537
 DB 2013 AGCCAGTCAAGAGAGAAACCATAGATCATAGATGTAATATATGATCATCTGGAACC 2072
 QY 538 -----AsnMetLeuArgLeuGlyLysSer 545
 DB 2073 CCTCAAAAGGCCCTGAACCCCTTTTGTGTAGCATATGCTGAGGCTTGGAATAATCA 2132
 QY 546 GluProTyrThrLeuAlaLeuGluAsnValIleGlyAlaLysAsnMetAsnValArgPro 565
 DB 2133 GAACCTGGACCCCTGACATTTGGAATTTGTAGAGC-AAGAACATCAATGATTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrIlePheLysAspGlnAsnLysAsnSerPhe 585

DB 2192 CTGCTCACTACTTATAGCCCTTATTTACTTGCTGCAAGACCAACAGATCTTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSe 605
 DB 2252 GTGGATGAGAGTACCGACATGAGATCCATATGACAGCCCAACATCAAGTGAAGATNAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuPhe 625
 DB 2312 CCTTAAATAGCTCTTGGAGATTAAGCATATGAAAGGAGCAAGTAAATGTACCTGTT 2371
 QY 625 eArgSerSerValAlaTyrAlaMetArgIleThrPheLysValLysAsnGluMetI 645
 DB 2372 CCGATATCTGTGCTATATGCTATGAGGAGTACTTTTAAAGTAAATCAATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPh 665
 DB 2432 TCTTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAACCAAGATCTCTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 DB 2492 CTTTGTCACTGCACCTTAAATGTGTGATATCATCTCTAGAACTGAAGTGAAGAGC 2551
 QY 685 aIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlu 705
 DB 2552 CATCAGAGATGCCCGAGCCGATTCATATGATGTTCCCTGATGACACAGCCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGluProProAsnGlnProPro 720
 DB 2612 GTTTCGGGATACAGCAACACTTGACCTCTTACAGGCCCT 2657
 RESULT 17
 US-10-143-114-71
 ; Sequence 71, Application US/10143114
 ; Publication No. US20030036180A1
 GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Geriltsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gurney, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P9330R1C211
 CURRENT APPLICATION NUMBER: US/10/143,114
 PRIOR APPLICATION DATE: 2002-05-09
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 71
 LENGTH: 3732
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-143-114-71
 Alignment Scores:
 Pred. No.: 0
 Score: 3720.00
 Percent Similarity: 83.88%
 Best local Similarity: 83.76%
 Query Match: 96.15%
 DB: 9
 Length: 3732
 Matches: 717
 Conservative: 1
 Mismatches: 2
 Indels: 138
 Gaps: 1

US-09-978-385-2_COPY_19_738 (1-720) x US-10-143-114-71 (1-3732)

QY	SeFTriIleGluGluGlnAlaIlySThrPheLeuAspTyrPheAsnHisGlnAlaGluAsp	20
Db	TCACACATGAGGACAGCCAGACATTTTGGACAGTTTAAACCAAGACCGAAGAC	15
QY	LeuPheTyrGlnSerSerIleuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsn	40
Db	CTGTTCATCAAAAGTTCACCTGGCTTGGAAATTTAAACCAACCAATATTAAGAAGAT	21
QY	ValGlnAsnIleAsnAsnAlaGlyAspIlySTPSPSerAlaPheLeuTyrGluGlnSerThr	60
Db	GTCCAAAACATGAATTAATGCTGGGGCAATGCTCTGCTTTTAAAGAAACAGTCACA	27
QY	LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValIlyLeuGlnLeuGln	80
Db	CTTCCCAAAATGATTCACATCAACAAAGAAATTCAGAAATTCACAGTCACCTCAGCTCAG	33
QY	AlaLeuGlnGlnAsnGlySerSerValIleuSerGlyAspIlySerTyrArgLeuAsnThr	100
Db	GCTTTACGCAAAATGGGCTTCAGGCTTCAGAAACAAAGACAAACGGSTTAACACA	39
QY	IleLeuAsnThrMetSerThrIleTyrSerThrGlyValIcyAsnProAsnPro	120
Db	ATTCTAAATCAATGAGCACACATCTCTACACTCGGAAAAGTTTGAATCCAGATTAATCCA	45
QY	GlnGluTyrLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr	140
Db	CAAAATAGCTATTACTCTGACACAGGTTTGAATGAATTAAGCAACAGTTTGAAGCTAC	51
QY	AsnGluArgLeuThrPalaTrpGlnSerTyrPargSerGlyValGlyGlnLeuArgPro	160
Db	AATGAGAGCGCTCGGGCTTGGGAAAGCTGAGATCTGAGTCGGGAACAGCTGAGGCCA	57
QY	LeuTyrGlnGluTyrValValLeuTyrAsnGluMetAlaArgAlaAsnHisTyrGluAsp	180
Db	TTATATCAAAAGATATGCTGTGTAAGAAATGAGATGGCAAGCAAAATCATTAAGAGAC	63
QY	TyrGlyAspTyrTrpArgGlyAspTyrGlyValaAsnGlyValaAspGlyTyrAspTyrSer	200
Db	TATGGGATTAATGGAGAGAGACATGAATTAAGGATGAATGGCTATGACTACAGC	69
QY	ArgGlyAlaLeuLeuLeuAspValaGluHisThrPheGluGluIleTyrProLeuTyrGlu	220
Db	CGGGCGACGTGATTAAGATGAGACATACCTTTGAAGAGATTAACCATTAATATGAA	75
QY	HisLeuHisAlaTyrValaArgAlaIlyLeuMetIleAsnAlaTyrProSerTyrIleSerPro	240
Db	CATCTTCATGCTATGATGAGGCAAAAGTTGATGATGCTTCTCTATACAGTCCA	81
QY	IleGlyCysLeuProAlaHisIleLeuGluTyrAspMetTyrGlyArgPheTyrPheAsnLeu	260
Db	ATTGGAGAGCTCCCTGCTCATTTGCTGGTGAATAGTGGGTAAATTTTGGCAAAATCTG	87
QY	TyrSerLeuThrValProPheGlyGlnTyrProAsnIleAspValaTrpAspAlaMetVal	280
Db	TACTCTTTGACAGTTCCTTTGGACAGAAACCAACATAGATGACTGACAAATGGTG	93
QY	AspGlnAlaTrpAspAlaGlnArgIlePheTyrGluAlaGluTyrPhePheValSerVal	300
Db	GACCAAGGCTGGGATGACACAGAAATTAATCAAGGAGCGCGAGAAATGTTTGTATCTGTT	99
QY	GlyLeuProAsnMetThrGlnIlyPheTrpGluAsnSerMetLeuThrAspProGlyAsn	320
Db	GGTCTTCCTAATATGACTCAAGAGATTTCTGGGAAATTCATGCTTACAGGACCCAGGAAAT	105
QY	ValGlnValaValaIcyHisPProThrAlaTrpAspLeuGlyIlyGlyAspPheArgIle	340
Db	GTTCACAAAGACAGTGGCCATCCACACCTTGGGACCTGGGGAAGGCGCACTCAGATAC	111
QY	LeuMetCysThrTyrValaThrMetIleAspPheLeuThrAlaHisAsnGluMetGluHis	360

Db	111	CTTATGTGCAAAAGGTGACATGGACGACTTCCTGACAGCTCATCATGCAATGGCGCAT	1173
QY	361	IIeGIINTyrAspMetAlaTyrAlaIaGIInProPheLeuLeuArgAsnGlyAlaAsnGlu	380
Db	1174	ATCCAGGTATGATATGGCATATAGCTCCACAAACCTTTTCTGTATGAAATAGAGCTATATGAA	1233
QY	381	GIYPheHISGlyAlaValGlyIuLeMetSerLeuSerAlaAlaThrProLysHISLeu	400
Db	1234	GGATTCCATGAAGCTGTTGGGAAATCATATGACTTTCTTGACGGCCACCTACACATTTTA	1293
QY	401	LysSerIIeGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe	420
Db	1294	AAATCCATATGGCTCTCTGCTCCACCCCATTTTCAAGAGACATGAAGACGAATATAACTTC	1353
QY	421	LeuLeuLysGlnAlaLeuThrIleValGlyIuThrLeuProPheTrpMetLeuGluLys	440
Db	1354	CTGCTCAACAGACACTCAGCATGTTGGGACTGCGCATTTACTTACATGTTAGAGAG	1413
QY	441	TrpArgTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrp	460
Db	1414	TGGAGGTGATGCTGTTTAAAGGGGAATTCGCAAGACGAGTGAGTGAATAAATGGTGG	1473
QY	461	GluMetLysArgGluIleValGlyAlaValGluProValProHISAspGluThrTyrCys	480
Db	1474	GAGATGAAGCGAGATAGATGGTGGGTGGAGACCTGGCCCATGATGAACAATACTGCT	1533
QY	481	AspProIaSerLeuPheHISValSerAsnAspTyrSerPheIleArgTyrTrpThrArg	500
Db	1534	GACCCCGCATGCTGTTCCATGTTTCTGATGATATACATTCATTCATTCATATACACAGG	1593
QY	501	ThrLeuTyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHISGluIlePro	520
Db	1594	ACCCCTTTCACAAATTCACATTTTCAAGAGACACTTGTCAAGACGCTTAAACATGAGGCCCT	1653
QY	521	LeuHISLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe	537
Db	1654	CTGCACAAATGTGACATCTCAAACTCTACAGAGACTGGACGAAACGTGT- GTTAGAAT	1712
QY	537	-----	537
Db	1713	ACCTCAAAATGTGAACTCTCTCTAGTATTCAGTATACATATTCATGCTCAGGTTTG	1772
QY	537	-----	537
Db	1773	TATTTGATTTCTTGTGTTCTAAAAAGAAAATTTATGCGCTCAAAAATGTCCTCATTTTCAAA	1832
QY	537	-----	537
Db	1833	ACCAAACTTTAATTTGTGTCAGACAGAGAACTAGACATACAAACAATTTGGGTGGGCCA	1892
QY	537	-----	537
Db	1893	CCCTCTTTCTCCCTATCATTAACAGACCCCTCTCTCTGTTAATTGGAAAGAAAGCGC	1952
QY	537	-----	537
Db	1953	GTTTAGGGGTGAATATATCTGTTAATATGCAATCTTTCTTATCTCGCAGAACCAATTT	2012
QY	537	-----	537
Db	2013	AGCCAACTCAAAAGAGAAAGAAACCATAGATCATAGATGTAATATATGATACATCTGAGACC	2072
QY	538	-----	538
Db	2073	CTCTCAAAAGGCCCTGGAACCCCTTTTTTTGTGTAGCAATATGCTGAGGCTTGGAAAAATCA	2132
QY	546	GluProTyrThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro	565
Db	2133	GAAACCTTGGAACCTCAGCATTTGAAAAAGTTGTGAGAGAC- AAGAACAATGAAATGTAAAGCCA	2191
QY	566	LeuLeuAsnTyrPheGluProLeuPheThrTyrLeuLysAspGlnAsnLysAsnSerPhe	585
Db	2192	CTGCTCACTACTTGTGAGCCCTTATTTTAACTGCTGTAAGAACCAAGAACGAATTTCTTTT	2251

QY	586	ValGlyIrrpserThrAspTppserProTyrAlaasp-GlnSerIleIysValArgIleSe	605
Db	2252	GTGGGAGGAGGTACCGACTGGAGTCATATGCAACCCAAAGCATCAAGTGAAGATAG	2311
QY	605	rIeuIysSerAlaLeuGIysApIysAlaTYrGIuTrPshAspAspAncIuMetTYrLeuPh	625
Db	2312	CCrTAATTCACCTCTTGAGAGTAAAGCATATNGATGGAGACAAATGAATGTACGTGT	2371
QY	625	eArgSerSerValAlaTYrAlaMetIArgGluTYrPheLeuIysValIysAsnGluMetI	645
Db	2372	CCGATCATCTGTTCATATGCTATGAGGACGACTTTTAAACGTAAAAAAATCAGATCAT	2431
QY	645	eIeuPheGlyIuGIuAspValaTYrAlaAsnLeuIysProArgIleSerPheAsnPh	665
Db	2432	TCCTTTTGGGGAGGAGGATGTGCCAGTGGCTAAATTGAACCAAGAACTCTCTTAAATT	2491
QY	665	ePheValThrIaIaProIysAsnValSerAspIleIeProArgThrGIuValGIuIysAl	685
Db	2492	CTTTGTCTCCTCACCTATAAATGTGTGTGATATCATTTCTTGAACCTGAAGTTGAAAAAGC	2551
QY	685	aIleArgMetSerArgSerArgIleAsnAspAlaPheArgIleuAsnAspAsnSerIeuGI	705
Db	2552	CATCAGAGATGTCCCGGAGCCCTATCAAAAGATGCTTCCGTGTGATGACAACAGCCTAGA	2611
QY	705	uPheLeuGIyIleGIuProThrLeuGIuProProAsnIuProPro	720
Db	2612	GTTTCTGGGGATACAGCCACACTTGGACCTCTTACCAAGCCCCCT	2657

QY 361 IleglnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
 DB 1174 ATCCAGTATGATATGGCATATGCTGCACACCTTTCTGCTAAAGAAATGAGACCTAATGA 1233
 QY 381 GlyPheHISgluAlaValGlyGluIleMetSerLeuSerAlaIleThrProlystIleu 400
 DB 1234 GGATTCACATGAAGCTGTGGGGAATCATGTCCTTTCTGACGCCACACCTTAAGCATTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhe 420
 DB 1294 AATCCATGGTCTCTCTGACCCGATTTTCAGAAAGCAATGAAGAAATAAACCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleLeuGluLys 440
 DB 1354 CTGCTCAACAGACCTACACATGTGTGGAGCTGCGCATTTACTTACATTTGTAGAGAG 1413
 QY 441 TrpArgTyrMetValPheLysGlyGluIleProLysAspGlnTyrMetLysLysTyrTrp 460
 DB 1414 TGGAGTGGATGGTCTTTAAAGGGGAATTCACAAAGACAGTGGATGAAAAAGTGGTGG 1473
 QY 461 GluMetLysArgGluIleValGlyValGluProValProHISAspGluThrTyrCys 480
 DB 1474 GAGATGAAGCAGAGATAGTGGGTGGTGGACCTGCGCCATGATGAACATACATGCT 1533
 QY 481 AspProLysSerLeuPheHISValSerAsnAspTyrSerPheIleArgTyrTyrThrArg 500
 DB 1534 GACCCCGATCTCTGCTTCATGTTCTTCATGATTCATTCATTCATTCATTCATTCATTC 1593
 QY 501 ThrLeuTyrGlnPheGlnPheGlnGluLeuLysGlnAlaIleLysHISgluIlePro 520
 DB 1594 ACCCTTACCAATTCACATTCACAGAGACCTTTGTCAAGAGAGTAAACATGAAGGCCCT 1653
 QY 521 LeuHISLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 537
 DB 1654 CTGCACAAATGTGACATCTCAACTACACAGACCTGACAGAAACHTGT- GTAAAGAAAT 1712
 QY 537 ----- 537
 DB 1713 ACTCAAAATGTGACCTCTCCATGATTCAGATTACATTCATTCATTCATTCATTCATTC 1772
 QY 537 ----- 537
 DB 1773 TATTTGATTTCTTTGTTCTTAAAGAAATTTATGGCTCAAAATGCTCATTTACAA 1832
 QY 537 ----- 537
 DB 1833 ACCAAACATTTAATTTGTGTGACAGACAGAACCTACACATCAACATTTGGTGGGCCA 1892
 QY 537 ----- 537
 DB 1893 CCTCTTTTCCCTATCATATACAGACCTCTCTCTCTGTAATTTGAAGAAAGAGCG 1952
 QY 537 ----- 537
 DB 1953 GTTTAGGGTGAATATATCTGTTAATATGATTCCTTTCTTATCTGCAGAGCAAAATTT 2012
 QY 537 ----- 537
 DB 2013 AGCCAGTCAAAAGAAAGAAACCATATAGATGTAATATATATGATCTGGAACC 2072
 QY 538 ----- 538
 DB 2073 CCTCAAAAGCCCTGAACCCCTTTTGTGTAGCAATATGCTGAGGCTTGGAAATATCA 2132
 QY 546 GluProTyrPheIleAlaLeuGluAsnValGlyAlaLysAsnMetLysValArgPro 565
 DB 2133 GAACCTGTGACCTTACATTTGAAATGTTGTAAGGC- AAGAAACATGATATTAAGGCCA 2191
 QY 566 LeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCAACACTTGTAGCCCTTATTTACTGCTGCAAGACAGACAGCAAGAAATCTTTT 2251
 QY 586 ValGlyTyrSerThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleLe 605

DB 2252 GTGGATGGAGTACCAGCTGAGTCCATATTCAGACCCAAAGCATCAAGTGGAGTAAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPAsnAspAsnGluMetTyrLeuPhe 625
 DB 2312 CCTAAATCAGCTCTTGGAGATTAAGCATATGATGGAAGACAGACATGAATGATCTGTT 2371
 QY 625 eArgSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetI 645
 DB 2372 CCGATCATCTGTGTCATATGCTATGAGGAGCTACTTTTAAAGTAAAGTAAATCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsn 665
 DB 2432 TCTTTTGGGAGAGAGATGTGGAGTGGTATTTGAAACCAAGATCTCTTTAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAl 685
 DB 2492 CTTGTCTACAGCTTAAATATGCTGATATCATTCATGAACTGAAGTGAAGG 2551
 QY 685 aileArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
 DB 2552 CATCAGATGTCCCGGAGCCGTATCATATGATGCTTCCGTGAATGACACAGCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGGATACAGCAACACTGTGACCTCTCTCAACAGCCCT 2657

RESULT 19

US-10-142-419-71

 Sequence 71, Application US/10142419
 Publication No. US20030044945A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: Deforge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Flivaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3330R1C244
 CURRENT APPLICATION NUMBER: US/10/142,419
 CURRENT FILING DATE: 2002-05-10
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 71
 LENGTH: 3732
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-142-419-71

Alignment Scores:

Pred. No.: 0 Length: 3732
 Score: 3720.00 Matches: 717
 Percent Similarity: 83.88% Conservative: 1
 Best Local Similarity: 83.76% Mismatches: 2
 Query Match: 96.15% Indels: 138
 DB: 9 Gaps: 1

US-09-978-385-2_copy_19_738 (1-720) x US-10-142-419-71 (1-3732)

QY 1 SerThrIleGlnGluGlnAlaLysThrPheLeuAspLysPheAsnHISgluAlaGluAsp 20

```

|||||
Db 94 TCACACCTTGAAGAACGCGCAAGACATTTTGGACAAAGTTAAACCGAAGCGGAGAC 153
Oy LeuphethrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGluAsn 40
Db 154 CTGCTTATCAAAAGTTACCTTCTCTTGGAAATTAATACACCAATATATCTGAAGAAAT 213
Oy 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThr 60
Db 214 GTCCAAAACATGATATATGCTGGGACAAATGGCTCTCTTTTAAAGAAACAGTCCACA 273
Oy 61 LeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
Db 274 CTGGCCCAAAATGATCCATCAACAGAAATTCAGATCTCACAGTCAAGCTTCAGCTGAC 333
Oy 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThr 100
Db 334 GCTCTTACGCAAAATGGGCTCTGCTGCTCTCAAGAACCAAGCAACGCTTGAACACA 393
Oy 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
Db 394 ATTTAAATATCAATGACACCATCTACAGTACTGAAAGTTGTAAACCCAGATAAACCA 453
Oy 121 GlnGlnCysLeuLeuLeuGlnProGlnLeuAsnGlnIleMetAlaAsnSerLeuAspTyr 140
Db 454 CAAGAAATGCTTATTAATGAAACAGGTTTGATGATGAATGAATGCAAAACAGTTAGACTAC 513
Oy 141 AsnGlnArgLeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgPro 160
Db 514 AATGAGAGCTCTGCGCTGGGGAAGCTGAGATCTAGGTCGGCAAGCAGCTGAGGCCA 573
Oy 161 LeuTyrGlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAsp 180
Db 574 TTATATGAGAGTATGTCGCTCTGAAAAATGATGAGTGGCAAGCAAAATCAATTATGAGAC 633
Oy 181 TyrGlnAspTyrTrpArgGlyAspTyrGlnValAlaAsnGlnValAspGlyTyrAspTyrSer 200
Db 634 TATGGGATTAATGAGAGAGACTATGAAATTAATGAGGATGAGTATGAGCTATGACAGC 693
Oy 201 ArgGlnLeuLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGln 220
Db 694 CGCGGCGCATGATGAAAGATGGAACATACCTTTGAGAGATTAACCATTTATATGAA 753
Oy 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerPro 240
Db 754 CATCTTCATGCTATGTCAGGCAAGTGTATGATGCTATCTCTATATATCAAGTCCA 813
Oy 241 IleGlyCysLeuProAlaHisLeuLeuGlnLysPheMetTrpGlyArgPheTrpThrAsnLeu 260
Db 814 AATGGATGCTCTCTGCTCATTTGCTTGGTATATGTCGGGTACATTTTGGACAAATCTG 873
Oy 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
Db 874 TACCTTTGACAGATTCCTTTGACAGAAACCAACATAGATGTACTGATGCAATGAGTG 933
Oy 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerVal 300
Db 934 GACCAAGGCTGGGATGACAGAGATATATCAAGAGGCGGAGAAAGTTCTTGTATCTGTT 993
Oy 301 GlyLeuProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsn 320
Db 994 GGTCTTCTATATGACCTCAAGAGATTCGGGAAATTCATGCTTACCGAGCCCGAGAAAT 1053
Oy 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIle 340
Db 1054 GTTCGAAAGCAGTCCATCCACACAGCTTGGGAGCTGGGGAAGGCGACTTCAGAGATC 1113
Oy 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHis 360
Db 1114 CTTATGTCACAAAGATGACAAATGACACTTCTGACAGCTCATCAAGATGGGGAT 1173
Oy 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGln 380
|||||

1174 ATCCAGTATGATATGCGATATGCTGCACACCTTTTCTGCTAAGAAATGAGACTATGAA 1233
Oy 381 GlyPheHisGlnAlaValGlyIleMetSerLeuSerAlaIleThrProLysHisLeu 400
Db 1234 GGATTCATGAAAGCTGTGGGAAATCATGATCTCATTTCTGACGACACCTAAGCATTTA 1293
Oy 401 LysSerIleGlyLeuLeuSerProAspPheGlnGlnLysAspAsnGlnThrGlnLeuAsn 420
Db 1294 AATTCATTTGGCTTCTGTCACCCGATTTTCAAGAGCAATGAAGAAACGAAATTAACCTC 1353
Oy 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuLys 440
Db 1354 CTGCTCAAAACAGACATCAGATGTGGGACTGTGCCATTTACTTACATGTTACAGAAAG 1413
Oy 441 TrpArgTrpMetValPheLysGlnGlnIleProLysAspGlnTrpMetLysLysTrpTrp 460
Db 1414 TGGAGGTGATGGTCTTTTAAAGGGAAATTCCAAAGACCATGATGATAAAGTGGTGG 1473
Oy 461 GlnMetLysArgGlnIleValGlyValValGlnIleProValProHisAspGlnThrTyrCys 480
Db 1474 GAGATGAACGAGAGATAGTGGGGTGGGAACTGTGCCCATGATGAAGAAACATATAGT 1533
Oy 481 AspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArg 500
Db 1534 GACCCCGCATCTCTGTCATGTTCTGATGATTAATCTCATTCATTCATGATATACACAAAG 1593
Oy 501 ThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyPro 520
Db 1594 ACCCTTACCAATTTCCAGTTTCAAGAAAGCACTTTGCAAGCAGCTAAACATGAAGCCCT 1653
Oy 521 LeuHisLysCysAspIleSerAsnSerThrGlnAlaGlyLysLeuPhe 537
Db 1654 CTGCACAAATGATACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTT -GTAAGAAT 1712
Oy 537 537 1712
Db 1713 ACCTCAAAATGTTGAACCTCTCTCTAGTATTAATCAATTCATTCATTCATTCATTCAT 1772
Oy 537 537 1772
Db 1773 TATTTGATTTCTTTGTTCTTAAAGAAATTTTATGCTCAAAATGCTCATTTACAA 1832
Oy 537 537 1832
Db 1833 ACCAAACATTTAATTTGTGTGTGACAGACAGAACTAGACCATACACAAATGGGTGGCCA 1892
Oy 537 537 1892
Db 1893 CCTCTTTTCTCCATCATTAATACAGACCCCTCTCTGTTAATTTGAAGAAAGAGCG 1952
Oy 537 537 1952
Db 1953 GTTTAGGGTGAATATATCTGTTAATATGATCATTTCTTTATCTGCAGAAACAATTT 2012
Oy 537 537 2012
Db 2013 AGCCAACTCAAGAAAGAAACCATATGATCATGATTAATATATGTCATCTGGAACC 2072
Oy 538 538 2072
Db 2073 CCTCAAAAGCCCTGAACCCCTTTTGTGTAGCAATATGCTGAGGCTTGAAGAAATCA 2132
Oy 546 GluProTrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetLysValArgPro 565
Db 2133 GAACCCCTGGACCTTAGATTTGAAATGTTGTAGGAGC -AAGAAACATGATGTAAAGGCCA 2191
Oy 566 LeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
Db 2192 CTGCTCACTACTTATGAGCCCTTATTTACTGCTGATGAAGACCAAGAAATTTCTTTT 2251
Oy 586 ValGlyTrpSerThrAspTrpSerProTyrAlaAsp -GlnSerIleLysValArgIle 605
Db 2252 GTGGGATGAGATACCGAGCTGAGTCCATATGACAGACCAAGCAATCAAGATGAGATTAAG 2311

```

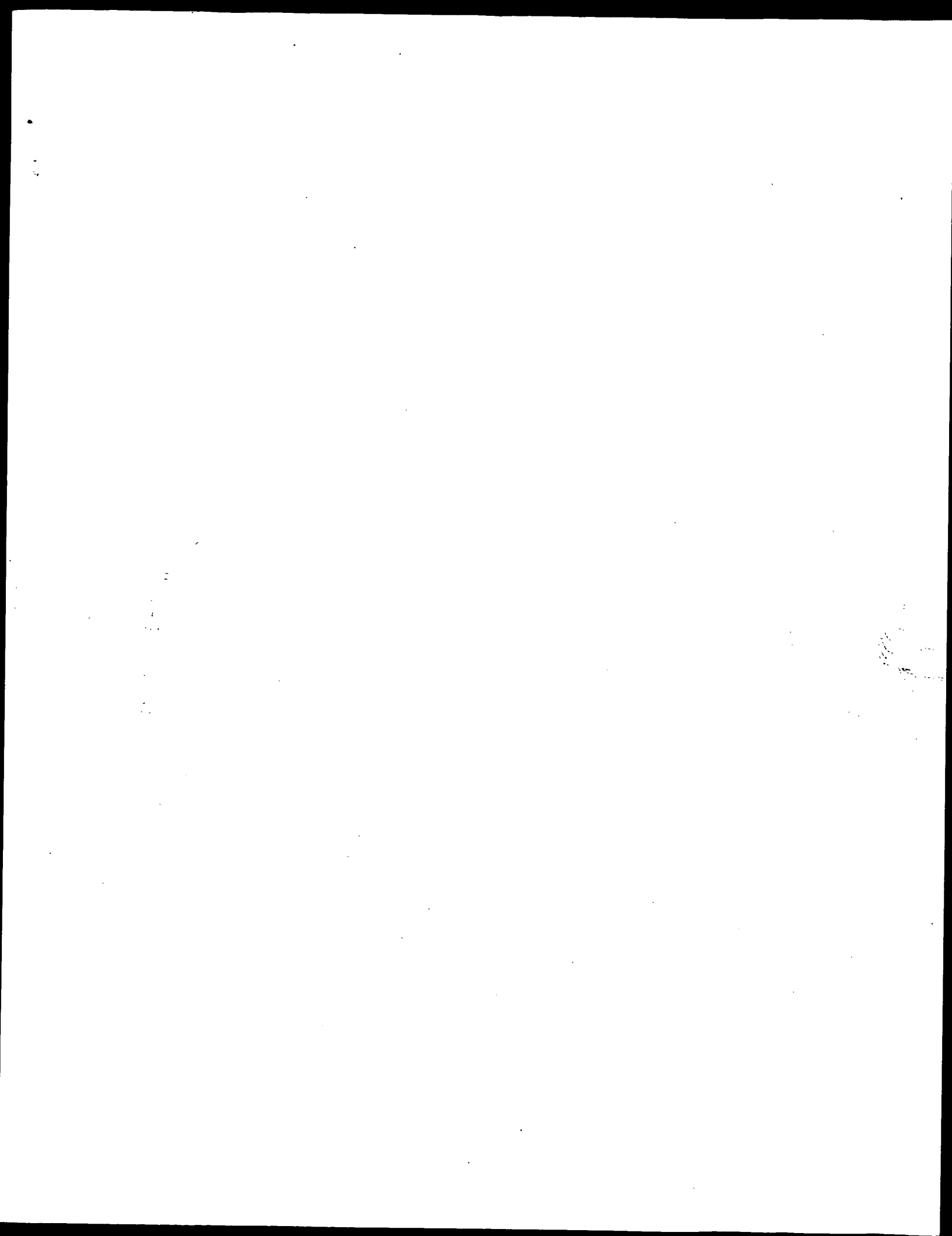
QY 605 rleuysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPh 625
DB 2312 CCTAAATACGCTCTGGAGATGAATGAATGAAGCAACGCAATGAATGACTCTGT 2371
QY 625 eArGSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetLi 645
DB 2372 CCGATCATGCTGTCATATGCTATAGGACGACTCTTTTAAAGTAAATAATCAATGAT 2431
QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLieserPheAsnPh 665
DB 2432 TCTTTTGGGAGAGAGATGCGAGTGCCTATTTGAACCAAGAACTCTTTAATTT 2491
QY 665 ePheValThrAlaProLysAsnValSerAspLleIleProArgThrGluValGluLysAl 685
DB 2492 CTTTCTACGACCTTAAATAATGTGCTGATATCTCTAGAACTGAAGTTGAAGAGGC 2551
QY 685 AlIeArgMetSerArgSerArgLLeAsnAspAlaPheArgLeuAsnAspAsnSerLeuG 705
DB 2552 CATGAGAGATGCCCGAGCCGATCAATGATGCTTCCGCTGTAATGCAACAGCTGAGA 2611
QY 705 uPheLeuGlyLleGlnProThrLeuGlyProProAsnGlnProPro 720
DB 2612 GTTCTGGGAGATACAGCAACACTTGACCTCTTAACCAAGCCCT 2657
RESULT 20
US-10-123-262-71
; Sequence 71, Application US/10123262
; Publication No. US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeBorge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C38
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-262-71
Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 3720.00 Matches: 717
Percent Similarity: 83.88% Conserves: 1
Best Local Similarity: 83.76% Mismatches: 2
Query Match: 96.15% Indels: 138
Gaps: 1
US-09-978-385-2_COPY_19_738 (1-720) x US-10-123-262-71 (1-3732)
QY 1 SerThrIleGluGluAlaLysTrpPheLeuAspLysPheAsnHisGluAlaGluAsp 20
DB 94 TCACACCATGTGAGAACAGGCCAAGACATTTTGGACAGATTAAACACGAGAGCCGAGAAC 153

QY 21 LeuPheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsn 40
DB 154 CTGTTCATCAAACTTCACTTCTTGGAAATTAATACCAATATTAATGAGAAAT 213
QY 41 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
DB 214 GTCCAAATGATATATGCTGGGACAAATGCTGCTTTTAAAGAACGCTCACA 273
QY 61 LeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB 274 CTGGCCAAATGATTCACATACAGAAATTCACAACTCACACTCAAGCTTCAGCTCAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
DB 334 GCTTTGACCAAAATGGCTTTCAGTCTCTCAGAAACAGACAGCAAGGTTGACACCA 393
QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnPro 120
DB 394 ATTCTAAATACAAATGAGCACCACTACAGTACGAAAGTTGTAAACCAATATGCA 453
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140
DB 454 CAAGATGCTTATTAATCTGAAACAGGTTGAATGAATATATGGCAAAAGTTTACAGCTAC 513
QY 141 AsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
DB 514 AATGAGAGGCTGGGCTTGGAAACCTGGAATCTGAGAGTGGCAAGCAAGCTGAGGCCA 573
QY 161 LeuTyrGluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAsp 180
DB 574 TTATATGAAAGATATGTGCTGTAATAATGAGANTGCAAGACCAATATATGAGGAC 633
QY 181 TyrGlyAspTyrTrpArgGlyAspTyrGluValAlaGlyValAlaAspGlyTyrAspTyrSer 200
DB 634 TATGGGATTAATGGAGAGAGACTATGAATGAATGGGTGATGCTATACATACAGC 693
QY 201 ArgGlyGlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlu 220
DB 694 CGCGGCCAGTTGATTAAGATGAGAACATACCTTTGAAGAAATTAACATTAATATGAA 753
QY 221 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLieserPro 240
DB 754 CATCTTCATGCTCATATGAGGCCAAGGTTGAGAAATGCTTATCTTCAATATACGTCCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeu 260
DB 814 ATTGATGCTCTCCCTGCTCATTTGCTGTGATATGTGGGTAGATTTGGCAAAATGTG 873
QY 261 TyrSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetVal 280
DB 874 TACTCTTGAAGATTCCTTGGACAGAAACCAATATGATGATGATGATGATGATGATG 933
QY 281 AspGlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
DB 934 GACCAAGCTGGGAGACAGAGAAATATCAAGAGAGGCCGAGAAATCTTTGATCTGT 993
QY 301 GlyLeuProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyLys 320
DB 994 GGTCTTCTTAATATGATCACTCAAGATTTCTGGAAATTCATGCTCAACGAGCCAGAAAT 1053
QY 321 ValGlnLysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgTyr 340
DB 1054 GTTCAGAAAGAGCTGTCATCCACAGCTTGGAGACCTGGGGAAGGAGGATTCAGATTC 1113
QY 341 LeuMetCysThrLysValThrMetLysAspPheLeuThrAlaHisHisGluMetCysHis 360
DB 1114 CTATATGTGCAAAAGGTGACATATGACACTTCTGACAGCTCATCATGATGAGGCGAT 1173
QY 361 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
DB 1174 ATCCAGTATGATATGCGATATGCTGCAACACTTTCTGCTAAGAAATGAGGATTAAGAA 1233

QY 381 GlyPheHisGluAlaValGlyIleMetSerLeuSerAlaAlaThrProLysHisLeu 400
 DB 1234 GGATTCATGAAGCTCTGGGAAATCAATGCTCTTTCGACCCACACTTAAGCTTTA 1293
 QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLeuAsn 420
 DB 1294 AATTCATGCTCTTCTGTCACCCGATTTTCAGAAAGACATGAACAAATAAATCTTC 1353
 QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrThrMetLeuGly 440
 DB 1354 CTGCTCAAAACACACACACAGATGTTGGACCTCCCATTTACTTACATGTTAGGAAG 1413
 QY 441 ThrAspTrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrp 460
 DB 1414 TCGAGTGGATGCTCTTTAAAGGGAATCCCAAGACAGCATGATGAAGAAAGTGCTG 1473
 QY 461 GluMetLysArgGluIleValGlyValValGluProValProHisAspGluThrTrp 480
 DB 1474 GAGATGAAGCGAGATAGTGGGGTGGTGAACCTGTCCTCATGTGAACATCTGT 1533
 QY 481 AspProLysSerLeuPheHisValSerAsnAspLysSerPheIleArgTrpThrArg 500
 DB 1534 GACCCGCACTCTGCTCCATGTTCTGATGATTTACTTCAATTCATTCATTCACACAG 1593
 QY 501 ThrLeuTrpGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyPro 520
 DB 1594 ACCCTTACCATTTCCAGTTTCAAGAAAGCACTTGTCAACGAGCTAAACATGAAGGCC 1653
 QY 521 LeuHisLysCysAspLysSerAsnSerThrGluAlaGlyGlnLysLeuPhe 537
 DB 1654 CTGCACAAATGTGACATCTCAAACTACAGAAAGCTGACAGAAACCTGTT-GRAGAAT 1712
 QY 537 537
 DB 1713 ACCCTCAAAATGTGAACCTCTCTAGATTTCAGATTACTTCCATTTCCATGCTAGTTG 1772
 QY 537 537
 DB 1773 TATTTGATTTCTTTGTTCTTAAAGAAATTTTATGGCTCAAAATGCTCATTTACAA 1832
 QY 537 537
 DB 1833 ACCAAACATTAATTTGTGTGACAGAGAACTAGACCATACAAATTTGGGTGGCCA 1892
 QY 537 537
 DB 1893 CCTCTTTTCTCCATCATTAATACTACAGCCCTCTCTCTCTGTAATTGGAAGAAAGCG 1952
 QY 537 537
 DB 1953 GTTTAGGCTGAAATATATCTGTTAATGATTTCTTTCTTATCTGCGAAGCAAAATTT 2012
 QY 537 537
 DB 2013 AGCCAAAGTCAAGAGAAAGAAACCATAGATCATAGATGAATATATGTCATCTGGAAC 2072
 QY 538 538
 DB 2073 CCTCAAAAGCCCTGAAACCCCTTTTGTGTAGCAATATGCTGAGGCTTGGAATAATCA 2132
 QY 546 GluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValArgPro 565
 DB 2133 GAACCTTGACCTGACCATTTGAAATGTTGAGAGC-AAAGACATGATGTAAGGCCA 2191
 QY 566 LeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPhe 585
 DB 2192 CTGCTCAACTTACTTGAAGCCCTTATTTACTGCTGGAAGACAGCAAGAAATCTTTT 2251
 QY 586 ValGlyTrpSerThrAspTrpSerProLysAlaAsp-GlnSerIleLysValArgIle 605
 DB 2252 GTGGGATGAGTACCGACTGAGTCCATATGACAGACCAAGCATCAAGTGAAGATTAAG 2311
 QY 605 rLeuLysSerAlaLeuGlyAspLysAlaTrpArgLysTrpAsnAspAsnGluMetLysLeu 625

DB 2312 CCTAAATCAGCTCTTGAGATTAAGCATATGAATGGAACGACATGAATGTAACCTGTT 2371
 QY 625 eArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysValLysAsnGlnMet 645
 DB 2372 CCGATCATCTGTTGATATGCTATGACGACACTTTTAAAGTAAATAATCAGATGAT 2431
 QY 645 eLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLysSerPheAsn 665
 DB 2432 TCTTTTGGGAGAGATGTCGAGTGAATTTGAACCAAGAAATCTCTTAAATTT 2491
 QY 665 ePheValThrAlaProLysAsnValSerAspIleLeuProArgThrGluValGlyVal 685
 DB 2492 CTTTGTCACTGCACCTTAAATGTTGCTGATATCATTTCTTGAAGCTGAAGTGAAGGC 2551
 QY 685 aLleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeu 705
 DB 2552 CATCAGAGTCTCCGAGGCGTATCATATGATGCTTCGCTGATGACCAACAGCCCTAGA 2611
 QY 705 uPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
 DB 2612 GTTCTGGGAGATACAGCCAACTTGACCTCTTAACCAAGCCCT 2657

Search completed: May 26, 2003, 20:28:02
 Job time : 349 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 17:59:14 ; Search time 2594 Seconds

(without alignments)
4495.278 Million cell updates/sec

Title: US-09-978-385-2_COPY_19_738
Perfect score: 3869
Sequence: 1 STRLEQAKTFIDKFNHEAD.....DNSLEFLGIQPTLGPNNPP 720

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xlh
-Q-cgna2_1/USPTO.spool/US09978385/runat.21052003.154420.19842/app.query.fasta.1.903
-DB-EST-QPMT-fastap -SUFFIX-rst -MINMATCH-0.1 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45
-DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-20 -MODE-LOCAL
-OUTFMT-pte -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09978385.ecgn.1.1.906.grunat.21052003.154420.19842 -NCPV-6 -ICPV-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEX-7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlum:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3827	98.9	3337	11 BC032938	BC032938 Homo sapi
2	1112	28.7	987	9 AL551235	AL551235 AL551235
3	1048.5	27.1	879	12 BG401683	BG401683 60246350
4	1031	26.6	766	13 BG962298	BG962298 602827090
5	975.5	25.2	763	13 B1913504	B1913504 603179133
6	940	24.3	676	13 B1561069	B1561069 603254275
7	924	23.9	635	12 BG722079	BG722079 602698682
8	857	22.2	555	9 AA162058	AA162058 ms31h11.r
9	829	21.4	1603	11 AK008530	AK008530 Mus muscu
10	824	21.3	609	13 B1561359	B1561359 602105037
11	813	21.0	965	12 BF789159	BF789159 602105037
12	781	20.2	452	12 BG428060	BG428060 602501471
13	749.5	19.4	556	13 BM030353	BM030353 488973 MA
14	741	19.2	686	10 BB652968	BB652968 BB652968
15	687.5	17.8	800	13 B1826471	B1826471 603076177
16	680	17.6	900	13 B1415293	B1415293 602987289
17	673	17.4	1018	13 BMS60537	BMS60537 AGENCOURT
18	661.5	17.1	848	13 B1904144	B1904144 603166751
19	648.5	16.8	741	12 BG772384	BG772384 602722306
20	646	16.7	848	13 B1766340	B1766340 603052444
21	625	16.2	348	9 AL039738	AL039738 DKEP434M
22	622	16.1	773	13 B1830936	B1830936 603081067
23	615	15.9	887	13 B1196733	B1196733 602755360
24	614	15.9	347	10 AM848579	AM848579 I13-CR021
25	593	15.3	666	13 B1686319	B1686319 60313252
26	588.5	15.2	725	13 B1821905	B1821905 603037059
27	581.5	15.0	842	12 BF784147	BF784147 602108008
28	578	14.9	751	13 B1828973	B1828973 603074849
29	564.5	14.6	823	13 B1914635	B1914635 603179402
30	564	14.6	606	10 AM322773	AM322773 uc02h11.y
31	558	14.4	492	10 AM260204	AM260204 um74f04.x
32	553	14.3	728	10 AM026259	AM026259 wv10g097.x
33	547	14.1	865	13 B1825221	B1825221 603071908
34	544	14.1	965	13 B1453005	B1453005 603170164
35	543.5	14.0	775	13 B1103238	B1103238 602889478
36	542	14.0	628	13 BM487251	BM487251 p0m2n.pk0
37	541.5	14.0	724	13 BM290186	BM290186 EST575720
38	538	13.9	749	10 AM026258	AM026258 wv10g097.x
39	535	13.8	896	12 BF783379	BF783379 602110217
40	534.5	13.8	969	12 BF780858	BF780858 602105748
41	533.5	13.8	755	9 A1971658	A1971658 w06g04.x
42	530.5	13.7	989	13 BM553966	BM553966 AGENCOURT
43	529.5	13.6	738	12 BF001772	BF001772 7933a10.x
44	527.5	13.6	723	12 BG644185	BG644185 ODC068 OI
45	527.5	13.6	954	14 BQ947556	BQ947556 AGENCOURT

ALIGNMENTS

RESULT 1
BC032938
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

BC032938 3337 bp mRNA
Homo sapiens, angiotensin I converting enzyme (peptidyl-di-peptidase A) 2, clone IMAGE:4830668, mRNA.
BC032938.1 GI:21432080
human.
HIT.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3337)
Strausberg,R.
Direct Submission

JOURNAL

Submitted (14-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: gcgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbiology.org

Anup Madan, Jessica Fahey, Erin Heltan, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://lmln.gov>
Series: IRAC Plate: 34 Row: 1 Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11225608
This clone has the following problem: frame shifted.

FEATURES

Source

1.3337
/organism="Homo sapiens"
/db_xref="locusID:59272"
/db_xref="taxon:9606"
/clone="IMAGE:4830668"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

BASE COUNT 1023 a 634 c 756 g 924 t
ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3337
Score: 3827.00 Matches: 717
Percent Similarity: 99.728 Conservative: 1
Best Local Similarity: 99.588 Mismatches: 2
Query Match: 98.918 Indels: 1
DB: 11 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BC032938 (1-3337)

QY 1 SerThrlleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
DB |||||||
94 TCCACCATGAGGAAAGGAGGCAAGCAATTTTGGACAAAGTTTAAACGACGAGCGAGAC 153
QY 21 LeuPheTYGlnSerSerLeuAlaSerTPAsnTYRAsnThrAsnIleThhGluGluAsn 40
DB |||||||
154 CTGTTCTATCAAGAGTTCACCTTCTCTTGGAATTTAAACCAATATTACTGAGAGAGAT 213
QY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTTPSerAlaPheLeuLysGluGlnSerThr 60
DB |||||||
214 GTCCAAAACATGAATATGCTGGGAGCAAAATGGCTGCTTTTAAAGAAAGACAGCCACA 273
QY 61 LeuAlaGlnMetTYRProLeuGlnLuiLleGlnAsnLeuThrValLysLeuGlnLeuGln 80
DB |||||||
274 CTTCGCCAAATGATATCCACTACAGAAATTCAGAAATTCACAGTCACAGCTCAGCTGAG 333
QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100
DB |||||||
334 GCTCTTCAGCAAAATGGTCTTCAGTGTCTCAGACAGCAAGCAAGCAAGGTTGACACA 393
QY 101 IleLeuAsnThrMetSerThrIleTYRSerThrGlyLysValCysAsnProAspAsnPro 120
DB |||||||
394 ATTCATAATACAAATGACGACCATCTACAGTACTGAAAAAGTTTAAACCCAGATTAACCA 453
QY 121 GlnGluCysLeuLeuLeuGluProGlyLeuAsnGlnLuiLleMetAlaAsnSerLeuAspTYR 140
DB |||||||
454 CAAAGATCTTATTACTTGAACAGGTTGAAATGAATGAATGAACAGATTGACTAC 513

QY 141 AsnGluArgLeuThrAlaTrpLysSerThrArgSerGluValGlyLysGlnLeuArgPro 160
DB |||||||
514 AATGAGAGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTCGCGCAACAGCTGAGGCCA 573
QY 161 LeuTYRGLuGluTYRValValLeuLysAsnGluMetAlaArgAlaAsnHisTYRGLuAsp 180
DB |||||||
574 TTATATGAAGAGTATGCTGCTTGAATAATGATGGCAAGCAAAATCATTAATGAGAGCA 633
QY 181 TYRGLYAspTYRTPArgGlyAspTYRGLuValAsnGlyValAspGlyTYRAspTYRSer 200
DB |||||||
634 TATGGGAGTATTTGAGAGAGAGACTATGAGTAATGGGAGTAATGGCTATGACTACAGC 693
QY 201 ArgGlyGlnLeuIleGlnAspValGluHisThrPheGluGluLuiLysProLeuTYRGLu 220
DB |||||||
694 CCGCGCCAGTTCATTTGAAGATGTAAGATACCTTTTGAAGAGATTAACCATTAATGAGA 753
QY 221 HisLeuHisAlaTYRValArgAlaLysLeuMetAsnAlaTYRProSerTYRLeuSerPro 240
DB |||||||
754 CATCTTCATGCTATGATGAGGCAAGTTGATGAAATGCTTCTCTTATATACATCA 813
QY 241 IleGlyCysLeuProAlaHisLeuLeuGlyAspMetTPGlyArgPheTPRAsnLeu 260
DB |||||||
814 ATTGATGCTCCCTGCTCATTTGCTGCTGATATGTGGGTAGATTTTGACAAATCTG 873
QY 261 TYRSerLeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetVal 280
DB |||||||
874 TACTCTTGACAGATTCCTCTTGGACAGAAACCAACATGATGTTACTGATGCAATGCTG 933
QY 281 AspGlnAlaTPRAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerVal 300
DB |||||||
934 GACCAAGGCTGAGGATGACAGAGATA-TTCAAGAGGCGCGAAGTTCTTGTATCTGTT 992
QY 301 GlyLeuProAsnMetThrGlnGlyPheTPRAsnSerMetLeuThrAspProGlyAsn 320
DB |||||||
993 GCTCTCTTATATGACCTCAAGATCTGGAATAATTCATGCTAACGAGCCAGCAAT 1052
QY 321 ValGlnLysAlaValCysHisProThrAlaTPRAspLeuLysGlyAspPheArgIle 340
DB |||||||
1053 GTTCAGAAAGAGCTGCTATCCACAGACCTTGGGAGGCGGCACTTCAGAGATC 1112
QY 341 LeuMetCysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHis 360
DB |||||||
1113 CTATATGTCACAAAGGAGCAATGAGACCTTCTGACACCTCATCATGATGAGGAGAT 1172
QY 361 IleGlnTYRAspMetAlaTYRAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 380
DB |||||||
1173 ATCCAGATATGATATGCGATATGCTGCCAACCTTTCTGCTAAGAAATGAGCTAATGGA 1232
QY 381 GlyPheHisGluAlaValGlyLuiLleMetSerLeuSerAlaAlaThrProLysHisLeu 400
DB |||||||
1233 GGATTCATGATAGCTGTTGGGAAATCATGCTTCTGACACCACTCAAGCAATTTA 1292
QY 401 LysSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLleAsnPro 420
DB |||||||
1293 AATTCATGATGCTTCTGTCACCCGATTTCAAGAAAGCAATGAACAAATAAATCTTC 1352
QY 421 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTYRMetLeuGluLys 440
DB |||||||
1353 CTGCTCAAAACAGACAGCATGTTGGGAGCTGCTCATTTACTTACATGTTAGGAAG 1412
QY 441 TPArgTPMetValPheLysGlyLuiLleProLysAspGlnTPRMetLysLysTPRTP 460
DB |||||||
1413 TGGAGGTGATGCTCTTTAAAGGGGAAATCCCAAGACAGTGATGAAAAAGTGGTGG 1472
QY 461 GluMetLysArgGluLuiLleValGlyValValGluProValProHisAspGluThrTYR 480
DB |||||||
1473 GATATGAAGAGAGATAGTTGGGAGTGGGAGGAACTGTCACCATGATGAACAAATCTGT 1532
QY 481 AspProAlaSerLeuPheHisValSerAsnAspTYRSerPheIleArgTYRThrArg 500
DB |||||||
1533 GACCCCGCATCTCTGTTTCATGTTCTATGATTTACTATTCATTCATATTCACAAAGG 1592

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 879)
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1335 row: d column: 13
High quality sequence stop: 619.
Location/Qualifiers
1. 879
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4594140"
/clone_lib="NIH_MGC_75"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site: 1;
Site1 (ggccgcctcgcc); Site: 2; Site2 (ggccatagcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCATATATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGCGGCGGACATG-3' (30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT 275 a 177 c 210 g 216 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.98e-117 Length: 879
Score: 1048.50 Matches: 218
Percent Similarity: 80.50% Conservative: 9
Best Local Similarity: 77.30% Mismatches: 12
Query Match: 27.10% Indels: 45
DB: 12 Gaps: 4

US-09-978-385-2_COPY_19_738 (1-720) x BG401683 (1-879)

OY 370 GlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyIle 389
DB 2 CAACCTTTTCTGTAGAAATGAGCATATGAAAGATTCATGAGCTGTGGGAAATC 61
OY 390 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuSerProAsp 409
DB 62 ATGCTACTTTCTGCAGCCACACCACTTAATCAATTCATGCTGTGTCACCCGAT 121
OY 410 PheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 429
DB 122 TTTCAAGACACATGAAACAGAAATTAATCTCTCTCAACACACCTCAGATGT- 180
OY 430 GlyThrLeuProPheThrTyrMetLeuGluLysTyrPArgTyrMetValPheLysGlyIle 449
DB 181 GGGACTCTGCATTTACTTACATGTTAGAAATGAGATGATGCTTTTAAAGGGGAA 240
OY 450 IleProLysAspGlnTyrMetLysLysTyrPArgTyrMetValPheLysGlyIle 469
DB 241 ATTCCCAAGACACGATGATGAAAGATGCTGTGAGATGAG- 282
OY 470 ValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSer 489
DB 282 ----- 282

OY 490 AsnAspTyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnIle 509
DB 283 -----ACCTTTACCAATTCACATTCACAGAA 309
OY 510 AlaleucysGlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSer 529
DB 310 GCATCTTGTCAACGACACTTAACATGAAAGCCCTCTGCACAAATGTGACATCTCAACACT 369
OY 530 ThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgGlyLysSerGluProTyrThr 549
DB 370 ACAGAGCTGAGACAGAAATGTTCAATATGCTGAGGCTTGAGAAATCGAACCTTGAGCC 429
OY 550 LeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyr 569
DB 430 CTAGCACTTGAAAGATGTTAGACCAAGAACATGATGTAAAGCCACTCTCACTAC 489
OY 570 PheGluProLeuPheThrTyrPheLysAspGlnAsnLysAsnSerPheValGlyTyrSer 589
DB 490 TTTGAGCCCTTATTTACCTGCTGTAAGACCAAGCAAGAAATCTTTTGTGGATGGAGT 549
OY 590 ThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLysSerAl 609
DB 550 ACCGACTGAGATCATATGACAGAACCAAGATCAAGATGAGATTAAGCTTAATCATGTC 609
OY 609 aleucGlyAspLysAlaTyrGluThrPAsnAspAsnGluMetTyrIleuPheArgSerVa 629
DB 610 T---GGAGATTAAGCA-TATGACTGAGC---ACATGAATGCTCTCTGCCGATCACTGT 662
OY 629 lAlaTyr-AlaMetArgIleTyrPheLeuLysValLysAsnGlnMetIlePheGlyG 649
DB 663 GCATATGTGATATTCAGAGATCACTTTTAAAGTAAACATGC-----ATGATCTTGGGG 716
OY 649 || 649
DB 717 AG 718

RESULT 4
LOCUS BG962298 766 bp mRNA EST 12-JUN-2001
DEFINITION 602827090P1 NCI_CGAP_CO24 Mus musculus cDNA IMAGE:4981606 5',
mRNA sequence.
ACCESSION BG962298
VERSION BG962298.1 GI:14349935
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 766)
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10983 row: d column: 23
High quality sequence stop: 708.
Location/Qualifiers
1. 766
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4981606"
/clone_lib="NCI_CGAP_CO24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.

FEATURES
source

Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCBI_CGAP Library."

BASE COUNT 233 a 156 c 179 g 198 t

Alignment Scores:

Pred. No.: 2.2e-115 Length: 766
Score: 1031.00 Matches: 210
Percent Similarity: 88.85% Conservative: 21
Best Local Similarity: 80.77% Mismatches: 23
Query Match: 26.65% Indels: 9
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BG962298 (1-766)

```

OY 398 LysHisLeuLysSerIleGlyLeuSerProAspPheGlnGlnAspGlnThrGlu 417
    |||
DB 2 AACCACTGTAATCCATGCTGCTTCGCCATCCGATTTCCAGAGATAGCGAACAGAG 61
    |||
OY 418 IleAspPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrMet 437
    |||
DB 62 ATAACTCTCTACTGAAACAGCGCATTTGGAACACTACCGTTTACTTACATG 121
    |||
OY 438 LeuGluLysTrpArgTrpMetValPheLysGlyIleProLysAspGlnTrpMet 457
    |||
DB 122 TTAGAGAAAGTGGAGTGGATGCTTCCGGGGTGAATTCCTCCAAAGAGAGTGGAGAA 181
    |||
OY 458 LysTrpTrpGlnMetLysArgGluIleValGlyValGlnProLysAspGln 477
    |||
DB 182 AACTGTGGAGATGAAAGCGGAGATCGGTGGTGGTGGAGCTCTGCTCATGATGAA 241
    |||
OY 478 ThrTyrCysAspProAlaSerLeuPheHisValSerAspTyrSerPheIleArgTyr 497
    |||
DB 242 ACATACGTCGACCTGATCTGCTGTCATGTTTAAAGATATACATTCATTCATTCAT 301
    |||
OY 498 TyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuGlnAlaValHis 517
    |||
DB 302 TACACAGGACCACTTACCAATTCGATTCAGAAAGCTTTGTCAGAGCTAAGTAT 361
    |||
OY 518 GlnGlyProLeuHisLysCysAspIleSerAspSerThrGlnAlaGlnLysLeuPhe 537
    |||
DB 362 AATGCTCTCTGCAAAATGTCATCTCAATTCACCTAAGCTGGCGAGAGTTGCTC 421
    |||
OY 538 AsnMetLeuArgLeuGlyLysSerGlnProTrpThrLeuAlaLeuGlnAsnValValGln 557
    |||
DB 422 AAGATGCTGAGTCTTGGAATTCAGAGCCCTGAGCCAAAGCCTTGGAAACATGCTGAGG 481
    |||
OY 557 ValAlaLysAsnMetLysValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLe 577
    |||
DB 482 AGCAAGAGAAATGATGATTAACACACGCTCAATTTCTCCAAACGTTGTTTGACTGGCT 541
    |||
OY 577 UlysAspGlnAsnLysAsnSerPhe-ValGly-TrpSerThrAspTrpSerProTyrAla 596
    |||
DB 542 GAAAGACAGACAGAAATCTTTCAGTGGGCGGAGAACCTGAAATGAGACCCATATGCC 601
    |||
OY 597 AspGlnSerIleLysValArgIle-SerLeuLysSerAlaLeuGlnLysPheLysAlaLysGln 616
    |||
DB 602 GACCAAGCAATTAACAGAGATTAAGCCCTAAGCTTCGAGCTGAGCTACATGATG 661
    |||
OY 616 LysTrpAsnAsp-AsnGlnMetTyrLeuPheArgSerSerValAlaTyrAlaMetArgGln 635
    |||
DB 662 ACTGAGACCAAGCAAGAAATGCTCTGCTGCTGATCATCTGT-GCATATGCCATGAGAAC- 719
    |||
OY 636 TyrPheLeuLysValLysAsnGlnMetIleLeuPheGlnGlnLysAsp 651
    |||
DB 720 GTCATGCTCATATCAAAAACAGACAGTCTTTCTA-GAGGAGGAT 766
    |||

```

RESULT 5 B1913504 763 bp mRNA linear EST 16-OCT-2001
LOCUS B1913504 603179133bp1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5243048 5'
DEFINITION mRNA sequence.
ACCESSION B1913504

VERSION B1913504.1 GI:16177911
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 763)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsb@remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LMNL1613 row: b column: 09
High quality sequence stop: 675.
Location/Qualifiers

FEATURES

1..763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5243048"
/clone_11b="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 237 a 167 c 170 g 189 t

Alignment Scores:

Pred. No.: 1.43e-108 Length: 763
Score: 975.50 Matches: 205
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 88.74% Mismatches: 12
Query Match: 25.21% Indels: 9
Gaps: 3

US-09-978-385-2_COPY_19_738 (1-720) x B1913504 (1-763)

```

OY 4 GlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuThr 23
    |||
DB 83 GAGAGACAGGCCAGAACATTTTGGACAAAGTTTAAACAGAGCCGAGAACCTGTTCTAT 142
    |||
OY 24 GlnSerSerLeuAlaSerTrpAsnTyrAsnTyrAsnIleThrGlnGlnLysValGlnAsn 43
    |||
DB 143 CAAAGTCACTTCTCTTGGAAATTTAAACACCAATATTTACGAGAGATGCCAATAAC 202
    |||
OY 44 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAlaGln 63
    |||
DB 203 ATGAATTAATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACACTTGGCCAA 262
    |||
OY 64 MetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 83
    |||
DB 263 ATGTATCCACTTAACAGAAATTCAGATTCACAGTCAAGCTTCAGGCTCTTCAG 322
    |||
OY 84 GlnAsnGlySerSerValLeuSerGlnLysAspLysSerLysArgLeuAsnThrIleLeuAsn 103
    |||
DB 323 CAAATGCGCTTCAGTCAAGTCTCAGAAAGCAAGCAAGCGTTGAACAATTTCAAT 382
    |||
OY 104 ThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGlnLys 123
    |||

```

Db 383 ACAATGAGCAGCATCTACACTGCGAAAAGTTTGAACCCAGATAATCCACAAGATGC 442
 QY 124 LeuleuLeugluProglyLeuAaungluileMetalaAaSerleuApyrAaungluA 143
 Db 443 TTAATCTGACACAGAGTTGACTGACATTAATGCAACAGTTTAACTATGAGAGG 502
 QY 144 LeutrpAlaThr---GluSerTPPArgSerGluValGlyLysGluLeuAArgProLeuTyr 162
 Db 503 CTCGGGGCTGGCGAAGAGCTGGACATCTGAGTGGCGAAGCAGCTGAGCCATTAATAT 562
 QY 163 GluGluTyrValValLeu-LysAaungluMetalaAaAaNHistyrGluAaPyrGlu 182
 Db 563 GAAGAGTATGTGCTCTTGAACATGATGAGCAAGCAATCATTAATGAGAGCTATG 622
 QY 182 YAspTYrTPArg-GlyAspTYrGluVal-AsnGlyVal-AspGlyTYrAspTYrSerA 201
 Db 623 GGATTTATCGGAGAGAGACTAATGAATGAGGATGAGTATGAGTATGAGTATGAG 682
 QY 201 gGly-GluLeuileGlu---AspValGluNHisthrPhe---GluGluileLysProLeuT 219
 Db 683 CGGACCAAGTTGATTCAGACAGACTGTGGACATACCTTTGACGAGACTTAACCCATAT 742
 QY 219 Yr-GluNHistLeuNHistAlaTyr 225
 Db 743 ACTGAACCTCTTCATGCTAT 763

RESULT 6

LOCUS B1561069 676 bp mRNA linear EST 05-SEP-2001
 DEFINITION 60324275F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296531 5',
 mRNA sequence.
 B1561069
 VERSION B1561069.1 GI:15448383
 KEYWORDS EST.

SOURCE

human.
 ORGANISM

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shlrahi
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11750 row: f column: 20
 High quality sequence stop: 674.

FEATURES

source

1..676
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5296531"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to 10^5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 215 a 134 c 161 g 166 t

Alignment Scores:

Pred. No.: 2 656-104 Length: 676
 Score: 940.00 Matches: 187
 Percent Similarity: 97.40% Conservative: 0
 Best Local Similarity: 97.40% Mismatches: 1
 Query Match: 24.30% Indels: 4
 DB: 13 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x B1561069 (1-676)

QY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAaPlyrPheAaNHistGluValGluA 20
 Db 109 TCCACCAATGAG 168
 QY 21 LeuPheTYrGlnSerSerLeuAlaSerTPPArgSerGluValGlyLysGluLeuA 40
 Db 169 CTGTTCTATCAAAAGTCACTGCTCTTGGAAATTAACCAATATATACGAAAGAAAT 228
 QY 41 ValGlnAaMetAaAaAaAlaGlyAspLysTPPArgSerAlaPheLeuLysGlu-GlnSer 60
 Db 229 GTCCAAAG 288
 QY 60 rLeuAlaGlnMetLysTPPArgSerValLeuSerGluAaPlyrSerLysArgLeuAa 80
 Db 289 ACTTGCCCAATGATATCCATCACTCAAGAAATCAAGATCTCACTCAAGATCTCACTCA 348
 QY 80 nAlaLeuGlnGlnAaNHistGlnSerValLeuSerGluAaPlyrSerLysArgLeuAa 100
 Db 349 GGCTCTTCAGAAAG 408
 QY 100 rIleLeuAaNHistMetSerThrIleTYrSerThrGlyLysValLysAaNHistProA 120
 Db 409 AATCTTAATCAATGAG 468
 QY 120 oGluGluCysLeuLeuLeuGluProGlyLeuAaNHistMetAlaAaSerLeuAaPly 140
 Db 469 ACAAG 528
 QY 140 rAsnGluArgLeuTPPArgSerGluValGlyLysGluLeuAaPlyrSerLysArgLeu 160
 Db 529 CAATGAG 588
 QY 160 oLeuTYrGluGluTyrValValLeu-LysAaNHistMetAlaAaNHistGluVal 179
 Db 589 ATTATATGAG 648
 QY 180 AspTYr-GlyAspTYrTPArgGlyAsp 188
 Db 649 GACTATGGGGGATTAATTCAG 676

RESULT 7

LOCUS BG722079 635 bp mRNA linear EST 08-MAY-2001
 DEFINITION 60269682F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4830668 5',
 mRNA sequence.
 BG722079
 VERSION BG722079.1 GI:14001266
 KEYWORDS EST.

SOURCE

human.
 ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 635)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shlrahi
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at:

http://image.llnl.gov

Plate: L1M10752 row: C column: 21

High quality sequence stop: 633.

FEATURES

source

1. 635

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4830668"
/clone_1ib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carinci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 202 a 131 c 144 g 158 t

ORIGIN

Alignment Scores:

Pred. No.: 2,2e-102 Length: 635
Score: 924.00 Matches: 180
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 23.88% Indels: 0
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x BG722079 (1-635)

QY 1 SerThrIleGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGluasp 20

Db 94 TCCACCAATGAGAACAGGCCAAGACATTTTG-GACAAAGTTTAAACACGAGCCGAGAGAC 152

QY 21 LeuPheTyrGlnSerSerLeuAlaSerTyrPheAsnThrAsnIleThrGluGluasp 40

Db 153 CTGTCTTATCAAAAGTTTCACTGCTTCTTGAATTTATACCAATATTTACTGAAGAGAT 212

QY 41 ValGlnAsnMetAsnAlaGlyAspLysTyrSerAlaPheLeuGluGlnSerThr 60

Db 213 GTCCAAACATGATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACA 272

QY 61 LeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

Db 273 CTTCGCCAAATGATCCATCAAGAAATTCAGATCTCAGTCAGTCAAGCTTCAGCTCAG 332

QY 81 AlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr 100

Db 333 GCTCTTCAAGCAAAATGGGTCTTTCAGTCTCTCAAGACAAAGCAACGTTGACACACA 392

QY 101 IleLeuAsnThrMetSerThrIleTyrSerThrIleLysValLysAsnProAspAsnPro 120

Db 393 ATTCTAATATCAAAATGAGACCATCTACAGTACGAAAGTTTGTACCCAGATTAATCCA 452

QY 121 GlnGluGlnLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyr 140

Db 453 CAAAGATCTTATTAATCTGAACCAAGCTTGAATGAATGAATGAACAGTATTAGACTAC 512

QY 141 AsnGluArgLeuTyrPalatPrpGluSerTyrArgSerGluValGlyLysGlnLeuArgPro 160

Db 513 AATGAGAGCGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTGGGCAACACCTAGAGCCA 572

QY 161 LeuTyrGlnGluTyrValIleLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluasp 180

Db 573 TTATTTGGAAGATATGTGTCTTGAATAATGATGAGCAAGACCAATCATTAATGAGAC 632

QY 181 Tyr 181

Db 633 TAT 635

RESULT 8

AA162058

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:374045
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 446.
Location/Qualifiers
1. 555
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:608613"
/clone_1ib="Stratagene mouse skin (#937313)"
/sex="females"
/tissue.type="whole skin"
/dev.stage="11 weeks old"
/lab_host="GOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Whole skin from 11 week old C57BL/6 female mce.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGACTTTTCTTTTCTTTTCTTTT 3'"

FEATURES

source

BASE COUNT 168 a 106 c 150 g 131 t

ORIGIN

Alignment Scores:

Pred. No.: 2.94e-94 Length: 555
Score: 857.00 Matches: 155
Percent Similarity: 89.73% Conservative: 11
Best Local Similarity: 83.78% Mismatches: 19
Query Match: 22.15% Indels: 1
Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AA162058 (1-555)

QY 108 IleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGluCysLeuLeuGln 127

Db 1 ATTATACAGTCTCTGAAAAAGTTTGCACCCAAAGAACCCACAAAGATCTTATTAATCTTGAG 60

QY 128 ProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPalatPrp 147

Db 61 CCAGGATTTGATGAATTAATGAGCGACAAACACA-GACTACAAAGCTGAGGCTCTGGCATGG 119

QY 148 GluSerTyrArgSerGluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValVal 167

KN0VPLEEDYRVSOLKRPSEFEVTSPOVSDVDPREVEDALRMSGRINDVFG
LNDNSLEFGLHPLEPPYOPVPTIMLIFGVYVALVYVGIILITVGIKGRKKNET
KRENSYDSMDIGKESNAGFONSDDAQSF"
polyA_signal
1584..1589
/note="putative"
polyA_site
1603
/note="putative"

BASE COUNT 465 a 321 c 366 g 450 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 4,95e-90 Length: 1603
Score: 829.00 Matches: 155
Percent Similarity: 88.83% Conservative: 20
Best Local Similarity: 78.68% Mismatches: 22
Query Match: 21.43% Indels: 0
DB: 11 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x AK008530 (1-1603)

OY 524 CyaspiPleseranserThrgluAlaGlyLysLeuPheasnmelleuAryleugly 543
DB 4 TGTGACATCTCAAAATTCAGTGAAGCTGGGAGAAAGTTGCTCAAGATGCTGAATCTTGGA 63
OY 544 LysSerGluProThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnVal 563
DB 64 AATTCAAGCCCTGGACCAAGCCTTGAAAGTGGTAGACCAAGGAATATGATGTA 123
OY 564 ArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAsnGlnAsnLysAsn 583
DB 124 AAACACCTGCTCAATTCACCAACCTGTTGACCTGAGGCTGAAAGAGCAAGCAAGAAAT 183
OY 584 SerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspGlnSerIleLysValArg 603
DB 184 TCTTTTGGGGTGAACACCTGATGAGGCCCAATGCCACCAACCAATTAAGTGAGG 243
OY 604 IleSerLeuLysSerAlaLeuGlyAspLysAlaTrpGluTrpAsnAspAsnGluMetTrp 623
DB 244 ATAAGCCTAAATAGCTCTTGAGCTAATGCAATGAAAGCAACCAAGAAATGTTTC 303
OY 624 LeuPheArgSerSerValAlaTrpAlaMetArgGlnTrpPheLeuLysValLysAsnGln 643
DB 304 CTGTCCGATCATCTGTGCAATGATGCCATGAGAAAGTAATTTTCAATCAATCAAAACACAG 363
OY 644 MetIleLeuPheGlyGluLysAspValArgValAlaAsnLeuLysProArgLysSerPhe 663
DB 364 ACAGTCTCTTTCTAGAGAGAGAGATGACGAGTGATTTGAAACCAAGAGTCTCCTTC 423
OY 664 AsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluValGlu 683
DB 424 TACTTCTTTGACCTCACCCCAAAATGTCTGATGATGATCTAGAGAGAGAGTTGAA 483
OY 684 LysAlaIleArgMetSerArgSerArgLysAsnAspAlaPheArgLeuAsnAspAsnSer 703
DB 484 GATGCCATCAGGATGTCGGGGCCGATCATATGATGCTTTGGCTGAAATGATTAACGC 543
OY 704 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProPro 720
DB 544 CTGAGGTTTCTGGGATTCACCCACACTGTAGCCACCTTACAGCCTCCT 594

RESULT 10
B1561359 609 bp mRNA linear EST 05-SEP-2001
LOCUS B1561359
DEFINITION 603255121F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5297380 5',
ACCESSION B1561359
VERSION B1561359.1 GI:15448673
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 609)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LAM11752 row: 1 column: 05
High quality sequence stop: 606.
Location/Qualifiers

FEATURES
source 1..609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5297380"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptPR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to 10⁶ 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 193 a 128 c 137 g 151 t
ORIGIN

Alignment Scores:

Pred. No.: 3.89e-90 Length: 609
Score: 824.00 Matches: 169
Percent Similarity: 97.70% Conservative: 1
Best Local Similarity: 97.13% Mismatches: 1
Query Match: 21.30% Indels: 4
DB: 13 Gaps: 0

US-09-978-385-2_COPY_19_738 (1-720) x B1561359 (1-609)

OY 1 SerThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHISGluAlaGluAsp 20
DB 94 TCCACCATTTGAGGAGAGGCGCAAGACATTTGTGACAAAGTTTAAACCAAGAGCGAAGAC 153
OY 21 LeuPheTrpGlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsn 40
DB 154 CTGTCTATCAAAAGTTCACCTGCTCTGGAATTAATTAACCAATTAATGAAAGAAAT 213
OY 41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerTh 60
DB 214 GTCCAAAACATGATATATGCTGGGGACAAATGTCTGCTTTTAAAGAGAAAGTCCAC 273
OY 60 rLeuAlaGluMetTrpProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuG 80
DB 274 ACTTGCACCAATGATATCCTACCAAGAAATTCGAATCTCAAGTCAAGCTTAAGCTGCA 333
OY 80 nAlaLeuGlnGln-AsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsn 100
DB 334 GGCTCTTACGACCAAAAGGGCTTCAATGCTCTCAAGAAAGCAAGAGCAAGGTTGAACA 393
OY 100 hr-IleLeuAsnThrMetSerThrIleTrpSerThrGlyLysValCysAsnProAspAsn 119
DB 394 CGAATCTTAATTAATGAGACACATCTCAAGTCTGAGAAAGTTGTAACCAAGATTAAT 453
OY 120 ProGlnGlyCysLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAsp 139
DB 454 CCACAGAAATCTTATTAATCTTGAACCAAGGTTTGAATTAATTAATGCAAAACAGTTTAGAC 513
OY 140 TyrAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArg 159


```

Db      514 TACATGAGAGAGCTGCTGGCTGGAGAAAGCTGAGATCTGAGTGGAGAG-CAGCTGAGG 572
Qy      160 ProleutyrglulgluValValleuLysAsnGlu 171
Db      573 CCATTATATGAGAGATGTGTGCTTGAATAAATGAG 608

RESULT 11
LOCUS   BF789159                      965 bp    mRNA    linear    EST 12-JAN-2001
DEFINITION 602105037F1 NCI CGAP_K1d14 Mus musculus cDNA clone IMAGE:4223312
VERSION   BF789159
KEYWORDS  BF789159.1 GI:12094195
SOURCE    EST.
          house mouse.
          Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ACCESSION BF789159
VERSION   BF789159
KEYWORDS  BF789159.1 GI:12094195
SOURCE    EST.
          house mouse.
          Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
AUTHORS   Contact: Robert Strausberg, Ph.D.
          Email: cgaabs-remail.nih.gov
          Tissue Procurement: jeffrey E. Green, M.D.
          CDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
          Cloned distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNLN at:
          http://image.llnl.gov
          Plate: LLM9811 row: 1 column: 09
          High quality sequence stop: 676.
          Location/Qualifiers
            1..965
              /organism="Mus musculus"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4223312"
              /clone_1lb="NCI-CGAP_K1d14"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: Kidney; Vector: pCMV-SPORT6; Site:1: NotI;
              Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.75 kb. Constructed by Life
              Technologies. Note: this is a NCI-CGAP library."
BASE COUNT 286 a 213 c 253 g 212 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.92e-88      Length: 965
Score: 813.00           Matches: 166
Percent Similarity: 79.82%      Mismatches: 16
Best Local Similarity: 72.81%      Indels: 43
Query Match: 21.01%           Gaps: 2
DB: 12

US-09-978-385-2_COPY_19_738 (1-720) x BF789159 (1-965)
Qy      1 SerThrIleGluGluGlnAlaIleValThrPheLeuAspLysPheAsnHisGluAlaGluAsp 20
Db      74 TCCCTCAGCAGGAGAAATGCCAGACATTTTAAACACTTATATCAGGAAGCTGAGAGC 133
Qy      21 LeuPheTyrglnSerSerLeuAlaSerTrpAsnTyrsAnthrAsnIleThrGluGluAsn 40
Db      134 CTGTCTATCAAGATGCTCTGCTTGGATATATATACATACTACTACTGAGAAAT 193
Qy      41 ValGlnAsnMetAnsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThr 60
Db      194 GCCCAAAAGATGAGCTGAGCCAGCCAAATGGCTGCTTTATGAGAGACAGCTAAG 253
Qy      61 LeuAlaGlnMetTyrrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGln 80

```

```

Db      254 ACTGCCAAAGTTTCTCATTACAGAAATCCAGATCCGATCATCAAGCTCAACTACG 313
Qy      81 AlaLeuGlnGlnAsnGlySerValLeuSerGluLysAspLysArgLeuAsnThr 100
Db      314 GCCCTTCAGCAAGATGGGTCTTCAGCATCTCAGCAGACAGACAAACAGTTGACACA 373
Qy      101 IleLeuAsnThrMetSerThrIleTyrrSerThrGlyLysValCysAsnProAspAsnPro 120
Db      374 ATTCTGAGACACCAATGAGACCACTTACACTGCGAAAGTTTGCACCAACCAACCCA 433
Qy      121 GlnGluCysLeuLeuLeuGluGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrr 140
Db      434 CAAGAATGCTTATTACTTACAGCCAGATTTGATGAATTAATGACACAAAGACAGACTAC 493
Qy      141 AsnGluArgLeuThrAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgPro 160
Db      494 AACTGTAGGCTCGGCGACATGAGAGGCTGAGGCTGAGAGTGGCAACACTGAGCGCG 553
Qy      161 LeuTyrgluGluTyrrValValleuLysAsnGluMetAla-ArgAlaAsnHisTyrgluAs 180
Db      554 TTGTATGAGAGAGATGTGTCTCTGAAAACGAGATGCGAAAGACAAACAAATTTAAACA 613
Qy      180 PTrpGlyAspTyrrTrpArgGlyAspTyrgluValAsnGlyValAspGlyTyrrAspTyrrSe 200
Db      614 CTATGGGGATTTATTTGAGAGGGGACTATGAAACAGAGAGACAGACAGATG-TCACTATTA 672
Qy      200 rArgGlyGlnLeuIleGlnAspValGluHisThrPheGluGluIleLysProLeuTyrrG1 220
Db      673 CCCGTACACGTATGAGAC---TGTAACGTACTTTCGCAAA--TCAGCATTTGGTGA 726
Qy      220 whLseuHISAlaTyrrValarg 227
Db      727 GCCTCTTCATGCCATGATGAGG 748

RESULT 12
LOCUS   BG428060                      452 bp    mRNA    linear    EST 14-MAR-2001
DEFINITION 602501471F1 NIH-MGC_75 Homo sapiens cDNA clone IMAGE:4615121 5'
          mRNA sequence.
ACCESSION BG428060
VERSION   BG428060.1 GI:13334566
KEYWORDS  EST.
SOURCE    human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
AUTHORS   Contact: Robert Strausberg, Ph.D.
          Email: cgaabs-remail.nih.gov
          Tissue Procurement: CLONTECH Laboratories, Inc.
          CDNA Library Preparation: CLONTECH Laboratories, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Cloned distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNLN at:
          http://image.llnl.gov
          Plate: LLM1367 row: n column: 18
          High quality sequence stop: 450.
          Location/Qualifiers
            1..452
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4615121"
              /clone_1lb="NIH-MGC_75"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: Kidney; Vector: pNR-LIB (Clontech); Site:1:
              SfiI (ggcgccctggcc); Site:2: SfiI (ggccatcagcc); 5' and
              3' adaptors were used in cloning as follows: 5' adaptor
              sequence: 5'-CACGGCAATTATGAGCC-3' and 3' adaptor sequence:
              5'-ATTCTAGAGCGCAGGCGCGCCACATG-dt(30)BN-3' (where B = A,

```


C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library.

BASE COUNT 144 a 81 c 103 g 124 t

ALIGNMENT SCORES:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	4,45e-85	452	150	0	0	0
Best Local Similarity:	781.00					
Query Match:	100.00%					
	20.19%					

US-09-978-385-2_COPY_19_738 (1-720) x BG428060 (1-452)

QY 552 LeuGluAspValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 571
 DB 2 TTGGAAGAAGTTTATGAGCAGCAAGAACATGATGAGGCGCTGCTCAACTTGGAG 61
 QY 572 ProLeuPheThrTyrLeuLysAspGlnAsnLysAsnSerPheValGlyTyrSerThrAsp 591
 DB 62 CCTTATTTACCTGGCTGGAAGACCAAGAACAAATCTTTGTGGATGAGATCCGAC 121
 QY 592 TrpSerProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGly 611
 DB 122 TGGAGTCCATATCAGACCAAGACATCAAGTATGAGTAAAGCTTAATAGCTCTTGG 181
 QY 612 AspLysAlaTyrGluTyrPheAsnAspAsnGluMetTyrLeuPheArgSerSerValAlaTyr 631
 DB 182 GATTAAGCATATGATGAGTGAAGACCAAGAACAAATGATGATGCTGCTGATCATCTG 241
 QY 632 AlaMetArgGlnTyrPheLeuLysValLysAsnGluMetIleLeuPheGlyGluGluAsp 651
 DB 242 GCTATGAGGAGCTCTTTTAAAAAGTAAAAATCAGATGATCTTTTGGGAGAGGAT 301
 QY 652 ValArgValAlaAsnLeuLysProArgIleSerPheAsnSerPheValIleAlaProLys 671
 DB 302 GTGGAGTGGCTAATTTGAAACCAAGAAATCTCTTAAATTTCTTGTGCACTGACCTAAA 361
 QY 672 AsnValSerAspIleIleProArgTyrGluValGlyLysAlaIleArgMetSerArgSer 691
 DB 362 AATGTCTCTGATATCATCTCTAGAACTGAAGTTGAAAGGCAATGAGATGCTCCGGAGC 421
 QY 692 ArgIleAsnAspAlaPheArgLeuAsnAsp 701
 DB 422 CGTATCAATGATGCTTCCGTCTGATGATGAC 451

RESULT 13 556 bp mRNA linear EST 05-NOV-2001
 LOCUS BM030353
 DEFINITION 488973 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BM030353
 VERSION BM030353.1 GI:16743923

KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 556)
 Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Caeas,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,
 G.L., Heston,M.P., Isagrd,W.W., Rohrer,G.A., Chitto-McCown,C.G.,
 Perle,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 JOURNAL libraries and construction of a gene index for cattle
 MEDLINE Genome Res. 11 (4), 626-630 (2001)
 COMMENT 21180013
 Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemell.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.

PCR PRIMERS

FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGTCAGCAGC
 Plate: 120 Row: B Column: 16
 Seq primer: ATTATGTCACATATGAG.
 Location/Qualifiers

FEATURES

1..556
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 2BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 library made from pooled tissue from testis, thymus,
 semitendinosus muscle, longissimus muscle, pancreas,
 adrenal, and endometrium."

BASE COUNT 178 a 111 c 145 g 122 t

ALIGNMENT SCORES:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	4.68e-81	556	142	18	25	1
Percent Similarity:	749.50					
Best Local Similarity:	86.02%					
Query Match:	76.34%					
	19.37%					

US-09-978-385-2_COPY_19_738 (1-720) x BM030353 (1-556)

QY 53 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnIleGlnAsn 72
 DB 2 GCCTTTTAAAGAAAGATCCCGATGCGCAAACTTACTGCTGGAAGAAATTTAGAT 61
 QY 73 LeuThrValLysLeuGlnIleGlnAlaLeuGlnIleAsnGlySerSerValLeuSerGlu 92
 DB 62 CTCACACTCAAGGCTAATTTGAAGGCTCTTCAAGCATGAGGAGCTCAGCTCGCGCA 121
 QY 93 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 112
 DB 122 GAGAAAGAGCAAAAGATTTGAACACAGATTTAAATTAATGAGACCACTTCAAGTCTGG 181
 QY 113 LysValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGlnIleProGlyLeuAsnGlu 132
 DB 182 AAAGTTTGGACCCA--AATACACAGAGAGTGGCTTGAACCTTGAACCAAGTTTATGATGC 238
 QY 133 IleMetAlaAsnSerLeuAspTyrAsnGluIleTyrPalaIleTyrPalaIleTyrPargSer 152
 DB 239 ATATGGAAGAACAAGACAGACATCAATTCGAGGCTCTGGGCTTGGGAAGCTGAGGCT 298
 QY 153 GluValGlyLysGlnLeuArgProLeuTyrGluGlnIleTyrValValLeuLysAsnGluMet 172
 DB 299 GAAGTTGGCAAGCAACAAAGCCATGTGAAGAAGATCTGTCTCTAGAAATGAGATG 358
 QY 173 AlaArgAlaAsnIleTyrGluAspTyrGlyAspTyrTyrPargLysAspTyrGluValAsn 192
 DB 359 GCAGAGCCCAACATTTATGAGACATGAGGAGGAGCTATTGAGAGGGATTATAGTGAGCT 418
 QY 193 GluValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGluIleThrPhe 212
 DB 419 GGGGAGGAGGACATGATGATGACAGCCGTCAGCAGTGTGTAAGAATGAGAACCCACCTTT 478
 QY 213 GluGluIleLysProLeuTyrGluIleLysIleValIleValIleValIleValIleMetAsn 232
 DB 479 GCAGAGATTAACCATTAATGAGAACCACTTCAATGCTTATGTGAGGCGCAAGCTTGATGAT 538

QY 233 AlatyProserTyrile 238
|||||
DB 539 ACCGACCTTCATATC 556

RESULT 14
BB652968

LOCUS
DEFINITION BB652968 RIKEN full-length enriched, adult male hippocampus Mus
musculus cDNA clone C630041D11 5', mRNA sequence.
BB652968
VERSION BB652968.1 GI:15402926
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 686)

REFERENCE
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Koda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyai,T.,
Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)

TITLE
JOURNAL
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y., and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
Ishii,Y., and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Funct. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES
SOURCE
Location/Qualifiers
1..686
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="C630041D11"
/clone_id="RIKEN full-length enriched, adult male
hippocampus"
/sex="male"
/tissue_type="hippocampus"
/dev_stage="adult"
/lab_host="SOLR"
/note="Site_1: XhoI; Site_2: BamHI; cDNA library was

BASE COUNT 211 a 151 c 169 g 151 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 7,41e-80
Score: 741.00
Percent Similarity: 80.30%
Best Local Similarity: 73.23%
Query Match: 19.15%
DB: 10 Gaps: 0
US-09-978-385-2_COPY_19_738 (1-720) x BB652968 (1-686)
QY 1 Serhrllgllugllgllalalysrphleuaspyspshasnhtslgualagluasp 20
||| |||||
DB 86 TCCTCCACCGGAGAAATGCCAAGACATTTTAAACCACTTATCAGNAGCTGAAGC 145
||| |||||
QY 21 LeuphetrglnserSerleuAlaserPAsnTyrAsnThrAsnThrlnlebrglnuasn 40
||| |||||
DB 146 CTGCTTATCAAGATTCACCTGCTTCTTGGAATTTATATCTAACTTGAAGAAAT 205
||| |||||
QY 41 ValGlnsmetksnslalaglyaspyspserlaphleuaspyspserlph 60
||| |||||
DB 206 GCCCAAAAGATGAGTGAGCTGCAGCCCAATGCTGCTTATGAGCAAGCTGTAAG 265
||| |||||
QY 61 LeuAlaglnmetYrProleuglnlueglnslleuAlasleuThrVallyslleugln 80
||| |||||
DB 266 ACAGCCCAAGATTCACCTACACCAAGAAATCCAGACCTCCATCAAGGCTGAACG 325
||| |||||
QY 81 AlaleuglnslansnglSerSerValleuSerGluaspyspserlarspser 100
||| |||||
DB 326 GCCCTTCAGCAAGATGCTGCTTACAGCAGCTGTCAGCAGACAGCAAGCAAGTGAAC 385
||| |||||
QY 101 IleleuAsnThrmetserThrleuYrserThrnglylsvalCysAsnProasp 120
||| |||||
DB 386 ATTCTGAACACCAATGACACCATTTACAGTACTGCAAAAGTTGCAACCAAGAACCA 445
||| |||||
QY 121 GlnGluCysleuLeuLeuGluProGlyLeuAsnGluIleMetAlasnsrleuasp 140
||| |||||
DB 446 CAAGAAATGCTTATTTACTTGAGCAGATTTGATGAATTAATGCAAGCAAGCAAGCTAC 505
||| |||||
QY 141 AsnGluArgleuThrAlarprgluserrTPargsergluValglyslleuArpr 160
||| |||||
DB 506 AACTCTAGGCTGCGGAGATGAGAGGCGCTGAGGCTGAGGCTGCAACAGCTGAGGCC 565
||| |||||
QY 161 LeuTyrGluGluTyrValValleuLysAsnGluMetAlaArgAlaAsnHslygluas 180
||| |||||
DB 566 GTGTATGAGAGATATGCGGGGNCCTGAAAAACAGAGAGGAGCAAGCAATATTAACGA 625
||| |||||
QY 180 pTyrGlyAspYrTyrArgGlyAspYrGluValasnglyValaspsly 196
||| |||||
DB 626 CTATGGGATTTATTTGANNAGGGGACTTGAAGCCNAGGGAGCAAGG 675
||| |||||

RESULT 15
BIR26471
LOCUS
DEFINITION BIR26471 800 bp mRNA linear EST 04-OCT-2001
603076177p1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167987 5',
mRNA sequence.
ACCESSION BIR26471
VERSION BIR26471.1 GI:15938021
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1417 row: 3 column: 20
High quality sequence stop: 797.
Location/Qualifiers
1. 800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5167987"
/clone_lib="NIH-MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH-MGC library."
BASE COUNT 169 a 244 c 230 g 157 t
ORIGIN
Alignment Scores:
Pred. No.: 3,58e-73 Length: 800
Score: 687.50 Matches: 128
Percent Similarity: 68.30% Conservative: 53
Best Local Similarity: 48.30% Mismatches: 81
Query Match: 17.7% Indels: 3
Gaps: 4
US-09-978-385-2_COPY_19_738 (1-720) x B1826471 (1-800)
QY 212 PheGluGluLeuLysProLeuTYRGIuHisLeuHisAlaTYRValAlaGlyLeuMet 231
Db 9 TTCAGAGAGCTGACGACCTCTACCTACCTGATGCTGACGCGCGGCGCTGCAC 68
QY 232 AsnAlaTYR--ProSerTYRLeuSerProIleGlyCysLeuProAlaHisLeuGly 250
Db 69 CGTACTACAGGGGCGCCGACATCACTGAGGGGCGCCATCTCTGCTACCTGCTGGG 128
QY 251 AspMetTYRglaTYRglaTYRglaTYRglaTYRglaTYRglaTYRglaTYRglaTYR 270
Db 129 AACATGTGGCGGCGACCTGCTCAACATCTATGCTGTGGTGGCCCTTCCCTTCAGCC 188
QY 271 ProAsnLeuAspValIleThrAspAlaMetValAspGlnAlaIleThrAspAlaGlyIlePhe 290
Db 189 CCTCTGATGACACCAAGAGGCTATGCTAAAGAGGAGGCTGAGCGCCAGAGAGATGTT 248
QY 291 LysGluAlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTYR 310
Db 249 AAGAGGCTGATGATTTCTTCACTCCCTGGGCGTGTGGCCGCTGCTGATGTTGG 308
QY 311 GluAsnSerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAla 330
Db 309 AACAGTGTGATGTGGAGAACCAACGACGCGGCGGAGGTGTCTGCCAGGCTCGGCC 368

QY 331 TrpAspLeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValIleMetAsp 349
Db 369 TGGGACTTCTACCAAGAGGAGCTCCGATCAAGAGTGCACACCGGAGACTTGGAG 428
QY 350 AspPheLeuThrAlaHisIleGluMetGlyHisIleGlnTYRAspMetAlaTYRAla 369
Db 429 GACCTGTGTGGGCGCCACAGAAATGGCGCACATTCAGTATTCATGAGTCAAAAGAC 488
QY 370 GlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyGlu 389
Db 489 TTACCTGTGGCTTGGAGGAGGAGGTGCCAACCCTGCTTCATGTAGGCGCATGGGAGCTG 548
QY 390 MetSerLeuSerAlaAlaIleThrProLysHisLysSerIleGlyLeuLeuSerProAsp 409
Db 549 GTAGCGCTGTAGTGTCTACGCGCCAGGACCTGCACATCTCAACCTGCTGAGAGTGGAG 608
QY 410 PheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 429
Db 609 GGTGGCAGCGAC---GAGCATGACATCAACTTTCGATGAAAGATGGCCCTTGACAAGATC 665
QY 430 GlyThrLeuProPheThrTYRMetLeuGluLysTYRParGlyTYRMetValPheGlyGlu 449
Db 666 GCCTTATCCCTCTGAGTACTCTGCTGATCATGAGTGGCGCTGAGGATTTGAT-GGAAGC 724
QY 450 IleProLysAspGlnTYRMetLysLysTYRProGluMetLysArgGluIleValGlyVal 469
Db 725 ATCCACCAAGAGACTATACCGAGAGTGTGAGACTCTCAGGCTGAAGTACAGAGGCT 784
QY 470 ValGluProValPro 474
Db 785 CTGCCCCAGTCC 799
RESULT 16
B1415293 900 bp mRNA linear EST 14-AUG-2001
DEFINITION 60287289F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5143199 5',
LOCUS B1415293
DEFINITION mRNA sequence.
ACCESSION B1415293.1 GI:15176216
VERSION B1415293.1 GI:15176216
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldi, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1353 row: a column: 24
High quality sequence start: 9
High quality sequence stop: 873.
Location/Qualifiers
1. 900
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5143199"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pTZ19-Lac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRV; 1st strand cDNA was prepared from mRNA obtained from pooled

583 n-----SerPheValGlyTrp--SerThrAspTyrSerPro 594

137 CCAGATCTGACGAATGTATGGCCACATCCCGAATATGACAGACCTGTATATGGCATGC

137 CCAGATCTGACGAATGTATGGCCACATCCCGAATATGACAGACCTGTATATGGCATGC


```

Db      446 AAGGAGACTATTAACCGAGAGTGGTGAAGCTCAGGCTGATGATCAGGCTGTGCCCC 505
QY      472 ProvalProhIsaapGluThrTrpCysAspProAlaSerLeuHisValSerAsnAp 491
Db      506 CCAAGTCCCAAGATCCCAAGGTGACTTGGACCCAGGCTCCAAAGTCCAGCTCTCTGCGAAC 565
QY      492 TySerHeHeLeaGlyTrpThrArgThrLeuTrpGlnPheGlnPheGlnValLeu 511
Db      566 GTCCCATTCAGTCAGTACTTGTACGCTTCATCATCCATCCAGTCCAGTCCAGGCGGT 625
QY      512 CysGlnAlaAlaLys-----HisGlnGlyProLeuHisLysCysAspLysSerAsnSer 529
Db      626 GGGGTTCGGGCAAGCCGGGCAACAGGGTCCCTCCACAGAGGTGACATCTTACCAATCC 685
QY      530 ThGluAlaGlyLysLeuPheAsnMetLeuArgLysSerGlnProTrpThr 549
Db      686 AAGGAGACGAGGAGAGTCTGCGGAGTGCATGAGCTGGCTGACAGTAAAGCCGCGCA 745
QY      550 LeuAlaLeuGlnAsnValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTrp 569
Db      746 GAGGCCATGAGCTGATCACAGGCGCCTTACATGTCAGCCTCCGATGATGATAC 805
QY      570 PheGluProLeuPheTrpLeuLysAspGlnAsnLys 582
Db      806 TTCAGCCACTGACAGATGGCTGCTCACCGAGAACAG 844

RESULT 19
Bg772384      741 bp      mRNA      linear      EST 15-MAY-2001
DEFINITION   602722306F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4839289 5',
VERSION      Bg772384
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens.
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
            Toshitsuki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10774 row: k column: 02
            High quality sequence stop: 581.
FEATURES     Location/Qualifiers
            1..741
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4839289"
            /clone_id="NIH_MGC_97"
            /lab_host="DH10B"
            /note="Organ: testis; Vector: pBluescript (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
            size-selected for average insert size 2.2 kb and
            normalized to R0T 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT   212 a      188 c      165 g      176 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      1,93e-68
Score:          648.50
Percent Similarity: 86.67%
Best Local Similarity: 83.89%
Query Match:    16.76%
DB:             12
Gaps:           4

US-09-978-385-2_COPY_19_738 (1-720) x Bg772384 (1-741)

QY      1 SerThrIleGlnGlnAlaLysTrpPheLeuAspLysPheAsnHisGluAlaLysP 20
Db      112 TCCACCATGAGAGACAGGCCAAGACATTTTGGACAGATTATACCGAGCCGAGAAC 171
QY      21 LeuPheTrpGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGlnLysn 40
Db      172 CTGTTCATCAAAAGTCACTGCTCTTGGATTTAACCAATATTACTGAAAGANT 231
QY      41 ValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnSerThr 60
Db      232 GTCCAAACATGATATATCTGGGACAAATGCTGCTTATTAACGAGACAGTCCACA 291
QY      61 LeuAlaGln-MetTrpProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGln 80
Db      292 CTGGCCCATATGATATCCACCTACAAAGAAATTCAGATATCCAGTCAAGTTCAGTCA 351
QY      80 nAlaLeuGlnGlnAsn---GlySerSerValLeuSerGlnLysSerLys-ArgLeu 99
Db      352 GGCTCTTCGACCAACACTGGGTCTTCACTGCTCAGACAGAACAGAACACGTTGA 411
QY      99 snThrIleLeu-AsnThr-MetSerThrIleLysSerThrGly-LysValCysAsnPro 118
Db      412 ACACATTTCTCAAAATACATGACATGACCATCTACAGTACGAGAACAGTCTGTAACCCAG 471
QY      118 sPAsnProGlnGlnLysLeuLeuLeuGlnProGlyLeuAsnGlu-IleMetAlaAsnSer 137
Db      472 ATAAACCAAGAAATGTTATTACTTGAACAGTTTGATGACATCACTACGCGAACAGT 531
QY      138 LeuAspTrpAsn---GluArgLeuTrpAlaTrpLysSer-TrpArgSerGlnValGly 156
Db      532 TTAGACTACCAACATGACAGGCTGCTGGCTGGGAAAGCTGAGAGTCT---CGAGGTGC 588
QY      156 sGlnLeuArg-----ProLeuTrpGln-GluTrpValValLeu 168
Db      589 TGCAGCAGCTGAGGCGCACTTATATGAAGAGATATGCGCTTT 632

RESULT 20
Bg766340      848 bp      mRNA      linear      EST 25-SEP-2001
DEFINITION   603052444F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5202083 5',
VERSION      Bg766340
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens.
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1506 row: g column: 12

```

High quality sequence stop: 840.

FEATURES
Location/Qualifiers
1..848

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5202083"
/clone_lib="NH_MGC_122"
/lab_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NH_MGC library."

BASE COUNT 184 a 249 c 247 g 168 t
ORIGIN

Alignment Scores:

Pred. No.: 4.93e-68 Length: 848
Score: 646.00 Matches: 129
Percent Similarity: 49.46% Conservative: 35
Best Local Similarity: 34.68% Mismatches: 95
Query Match: 16.70% Indels: 93
Gaps: 5

US-09-978-385-2_COPY_19_738 (1-720) x B1766340 (1-848)

QY 105 MetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlyCysLeu 124
Db 7 ATGGAACACCACTACAGCGCTGGCCACTGTGTCCACCCGCAATGCG-----ACCTCCCTG 60
QY- 125 LeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlnGlyLeu 144
Db 61 CAGCTCAGCCCAATCTGACGAAATGTGATGCCACATCCCGAAATATGAGACCTGTTA 120
QY 145 TrpAlaTrpGlnSerTrpArgSerGluValAlGlyLysGlnLeuArgProLeuTyrGlnGlu 164
Db 121 TGGGCAATGGAGGGCTGGCAGACACAGCGGGGAGAGCCATCCTCCAGTTTAAACCGGAA 180
QY 165 TyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGlyAspTyr 184
Db 181 TACGTGAACTCAACCAAGCGCTGCCGCTCATGTGATGATGCAAGGGGACTCG 240
QY 185 TrpArgGlyAspTyrGlnValAlaAsnGlyValAlaSerTyrAspTyrSerArgGlyLeu 204
Db 241 TGGAGGTCTATGTACGAGACACATCCCTGGAG----- 273
QY 205 IleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeuHisAla 224
Db 274 ---CAAGACCTGAGCGGCTCTTCAGAGAGCTGCAGCCACTCTACCTCAACCTGCATGCC 330
QY 225 TyrValArgAlaLysLeuMetAlaTyr---ProSerTyrIleSerProIleGlyCys 243
Db 331 TACGTGGCGCCGCGCTCCACCGTCACTACGCGGCCGCCACATCAACCTGGAGGGGCC 390
QY 244 LeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTrpHisAsnLeuTyrSerLeu 263
Db 391 ATTCTGCTCACTGCTGGGACATGTGGCGGCGAGACCTGTCCAAACATCTATGACTTG 450
QY 264 ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 283
Db 451 GTGGTGGCCCTTCCTTACGCCCTCGATGACACACACAGAGGCTATGCTAAAGCAGGGC 510
QY 284 TrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeuPro 303
Db 511 TGGAGCCCGAGGATGTTTAAGAGGCTGTGATGATTTCTTCACTCCCTGGGGCTGCTG 570
QY 304 AsnMetThrGlnGlyPheTrpGlnLysAsnSerMetLeuThrAspProGlyAsnValGlnLys 323
Db 304 AsnMetThrGlnGlyPheTrpGlnLysAsnSerMetLeuThrAspProGlyAsnValGlnLys 323

Db 571 CCGGTGCCCTCCTGAGTTCTGGAACAAGTGCATGCTGGAGAGCCAGCGGGCGGGAG 630
QY 324 AlaValCysHisProThrAlaTrpAspLeuLysGlyAspPheArgIleLeuMetCys 343
Db 631 GTGTCGTGCCACAGGCTGGCTGGGAC----- 657
QY 344 ThrIysValIleThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGlnTyr 363
Db 657 ----- 657
QY 364 AspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHis 383
Db 658 -----TTCCTAC 663
QY 384 GlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIle 403
Db 663 ----- 663
QY 404 GlyLeuLeuSerProAspPheGlnGlnLysAsnGlnIleAsnPheLeuLys 423
Db 664 -----AACGACATGACATCACTTCTCATGAG 693
QY 424 GlnAlaLeuThrIleValAlGlyThrLeuProPheThrTyr-MetLeuGlnLysTrpArgTr 443
Db 694 ATGGCCCTTGACAAAGATGCGCTTATCCCTTACCTACCTCGTCGATCAGTGGCGCTG 753
QY 443 PheValPheLysGlyGlnIleProLysAspIleTrpMetLysLysTrpTrpGlnMetLys 463
Db 754 GAGGTAATTTGATGGAAGCATCACCAGAGAACTATTAACCAAGAGGTGTGAGCCCTGAG 813
QY 463 SarGlnIleValAlGlyValAlaGluProValPro 474
Db 814 GCTGAAGTACAGGGGTCTCTGCCCTCCCGCACTGCC 847

Search completed: May 26, 2003, 20:20:28
Job time : 2609 secs

